**Doctoral thesis submitted to**
**the Faculty of Behavioural and Cultural Studies**
**Heidelberg University**
**in partial fulfillment of the requirements of the degree of**
**Doctor of Philosophy (Dr. phil.)**
**in Psychology**

Title of the publication-based thesis

*Good Thoughts, Bad Thoughts?*
*Investigating the Nature of the Wandering Mind and How to Capture It*

presented by
Lena Steindorf

year of submission
2020

Dean: Prof. Dr. Dirk Hagemann
Advisor: Prof. Dr. Jan Rummel

**Table of Contents**

**Acknowledgements**

The past few years have been an exciting, challenging, frustrating, joyful, and most of all educational experience. Although I was supposed to focus on my research, my mind probably spent half of the time wandering. However, I have learned that this is not exclusively a bad thing. When writing a dissertation, a drifting mind can be incredibly annoying, but it also allows you to take a fresh look at pending problems, to have a short mental rest, or to wander towards pleasant memories, such as enjoyable experiences with other people. Indeed, research shows that mind wandering is oftentimes of a social nature, just as is the time spent as a doctoral student. Colleagues, friends and family surrounded me, and I am thankful for their support.

First, I would like to thank my supervisor Prof. Dr. Jan Rummel. I can count myself lucky to be a part of your *Gang*. Thank you for always taking the time to discuss research and to give helpful advice not only on my scientific writing, but also on teaching and early career issues, which were on my mind. I very much enjoyed investigating the wandering mind with your guidance.

I am also grateful to Prof. Dr. Klaus Fiedler, who immediately agreed to review this dissertation and who included our research group into his adaptive cognition colloquium as well as his Christmas parties.

Having enthusiastic student research assistants around was extremely helpful when it came to data collection. I would still be sitting in the lab collecting data if it were not for you. Equally important, I am thankful for the positive and lighthearted work atmosphere you established. A special thank you goes to Holly for proofreading a lot of my English writing, including this dissertation, throughout the years. I am also grateful to my other proofreaders Sebastian P. and Kathy.

The former members of Jan Rummel's Gang, Ivan and Sebastian D., as well as the rest of the pavilion "inhabitants", which includes Bianca, Katja, Lena, Ruth, Sabine, and Cornelia, made my time in the smallest building of the institute very special. It was always a great pleasure to chat with you, to discuss research, to go out for ice cream breaks, to watch hockey, to play minigolf, and most of all to just know that there were people around who were going through a similar phase in life. Thank you for creating this wonderful, supportive environment within our little building!

Sometimes I even left the pavilion. Having chicken or tofu curry, couscous, or a fancy salad with Veronika, Alica, Ulf, Stefan, Mischa, Marie, Annika, Nils, and Max has been an integral part of my week. Thank you for all those pleasant lunch breaks and conversations.

Outside of Heidelberg University, the *Mannheimer Gang*, especially Annki, Helen, Katharina, and Nadine, has been supporting me. I cannot put into words how thankful I am for your friendship and constant encouragement. A big thank you also goes to my old school friends, who offered much needed distraction from my (sometimes too) psychological lifestyle, and to Sebastian, who particularly supported me and endured all my moods during the final writing period. Thank you for always believing in me. Moreover, I would like to thank my family for always enabling me to pursue all my interests. Thank you for your continuous encouragement.

And finally, my grandmother has probably been my biggest supporter and constantly encouraged me to aim high. She was just as excited as I was when I started my time as a doctoral student, and I am sure that she would have been extremely happy and proud learning that I finally finished writing my dissertation. Grandma, I cannot thank you enough for your unconditional support.

## List of Scientific Publications of the Publication-Based Dissertation

### Manuscript 1

Steindorf, L., & Rummel, J. (2020). Do your eyes give you away? A validation study of eye-movement measures used as indicators for mindless reading. *Behavior Research Methods, 52*(1), 162-176. https://doi.org/10.3758/s13428-019-01214-4

### Manuscript 2

Steindorf, L., & Rummel, J. (2017). "I should not forget the apples!" – Mind-wandering episodes used as opportunities for rehearsal in an interrupted recall paradigm. *Applied Cognitive Psychology, 31*(4), 424-430. https://doi.org/10.1002/acp.3328

### Manuscript 3

Steindorf, L., Boywitt, C. D., & Rummel, J. (2020). *A fresh look at the Unconscious Thought Effect: Using mind-wandering measures to reveal thought processes in decision problems with high information load.* Manuscript submitted for publication.

### Manuscript 4

Steindorf, L., Hammerton, H. A., & Rummel, J. (2020). *Mind wandering outside the box – About the role of off-task thoughts and their assessment during creative incubation.* Manuscript submitted for publication.

**Good Thoughts, Bad Thoughts?**

**Investigating the Nature of the Wandering Mind and How to Capture It**

## Preface

*"Do penguins have knees? And how come their feet never get cold? Oh, where was I?"* Even though I am eager to write a preface to my dissertation and even though I find its topic particularly interesting, it is impossible for me to keep my thoughts from trailing off every once in a while. Sometimes, such *mind-wandering* episodes contain random content and my thoughts might drift towards the anatomy of penguins. Other times, I might try to figure out a solution to a personal problem or plan tonight's dinner, which many would probably consider more useful than thoughts about penguin knees. Independent of its content, the phenomenon that thoughts trail off from a task at hand is ubiquitous in everyday life. It is found that people spend up to 50 percent of their waking hours mentally occupied with thoughts that are unrelated to current external events (Killingsworth & Gilbert, 2010; Klinger, 1999), a fact that already demonstrates the high relevance and the resulting need for empirical investigations of the mind-wandering construct.

It probably comes as no surprise that mind wandering is often associated with negative outcomes, such as performance decrements within various domains (Mooneyham & Schooler, 2013; Smallwood & Andrews-Hanna, 2013). Everybody knows the feeling of being annoyed by drifting thoughts as they hamper focused work. Taking an important exam, reading a book, or talking to one's mother on the phone: The number of situations, in which mind wandering depicts a potential source of disturbance and error, is infinite. In some circumstances, it can even endanger others or ourselves, for example when driving at a high speed on the highway (Albert et al., 2018). Because of such far-reaching and omnipresent consequences, the costs of mind wandering have been extensively studied in the last two decades. The benefits, however, have initially been neglected. I believe that the phenomenon deserves to be viewed from a more balanced perspective. And indeed, current research is catching up on the positive side of mind wandering. It is argued that drifting thoughts could represent a source of inspiration and creativity, an opportunity for future planning or memory maintenance, a restful mental break, or a short and joyful escape from a currently unpleasant reality (Schooler et al., 2011).

This dissertation aims at painting a balanced picture of the wandering mind's nature, strengthening an adaptive view of the phenomenon. Figure 1 depicts a flow chart including the dissertation's main topics. First, negative consequences of drifting thoughts will be contrasted with positive ones, and I further aim to bring both sides of the medal more in line by focusing on thought-regulation processes. Previous insights will then be enriched by new contributions: I will introduce *memory* as a newly considered domain, which I found to benefit from mind-wandering processes. From a more methodological perspective and within the domains of *creativity* and *problem-solving*, I will not only take a look at possible further mind-wandering benefits, but also present a new research paradigm. This paradigm allows for the closer investigation of possibly thought-altering and intrusive effects of *thought probes*, the most widely applied mind-wandering assessment method. Thought probes ask participants about their current thoughts during a task and might thus modify the mind-wandering experience itself, thereby complicating the search for positive effects of the phenomenon. As such probes further rely on self-reports and thus contain a subjective component, I will finally report a review and validation study of eye-movement measures as objective mind-wandering indicators. Thus, the following dissertation presents an investigation of the nature of the wandering mind as well as of subjective and objective thought-assessment methods.

Oh, by the way, penguins do have knees. I googled it when my wandering mind needed to know the answer.

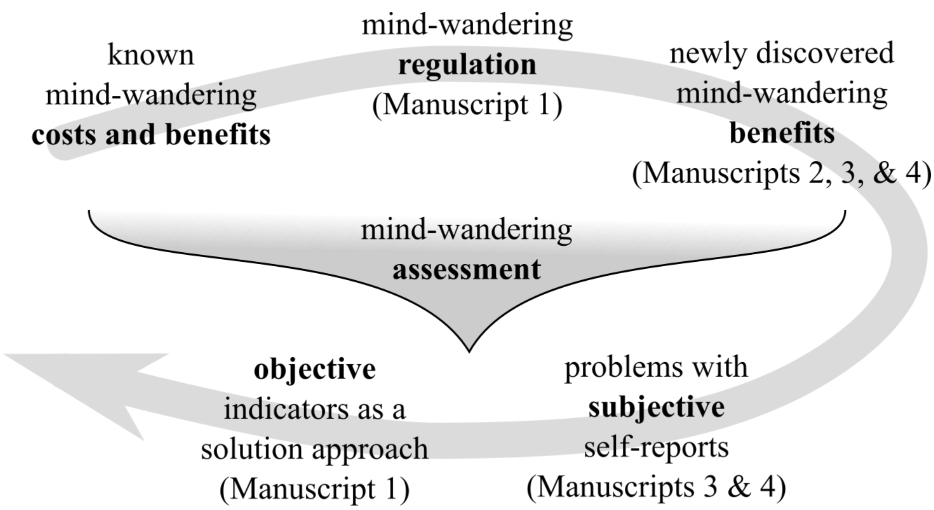


*Figure 1*. Flow chart. The figure integrates this dissertation's main topics, their order of consideration within the present work, as well as the respective manuscripts (in parentheses) covering the individual topics. When examining contentual topics such as costs and benefits, research results and conclusions are always dependent upon the quality of the employed assessment methods, which is why mind-wandering assessment is depicted in a curly bracket. This dissertation covers contentual as well as methodological topics concerning the mind-wandering concept.

# 1 Introduction to Mind Wandering

## *1.1 Definition*

Many everyday tasks require focused and sustained attention with so-called *task-related thoughts* enabling successful task-goal fulfillment, such as understanding the plot of a movie or solving a mathematical equation. However, thoughts sometimes trail off from the task at hand (Schooler et al., 2011). *Task-unrelated thought* (TUT) represents an attention shift, oftentimes towards personal issues, such as worries or unsolved problems (Mooneyham & Schooler, 2013). However, in general, people are able to generate task-unrelated content on an unlimited number of topics.

A description of mind wandering as TUT already represents one of many attempts at defining the construct (Seli et al., 2018). The TUT definition always involves an ongoing task and implies that the mind wanders as soon as thoughts are unrelated to this specific task, independent of the exact thought content or nature (Smallwood & Schooler, 2006). Other definitions cover, for example, thought in the absence of a current task, or the intentionality of thought, as terms such as *daydreaming* (Singer, 1966; Singer & Antrobus, 1963), *stimulus-independent thought* (Teasdale et al., 1995), *unintentional thought* (McVay & Kane, 2010), or *meandering* and *unguided thought* (Christoff et al., 2016) and their underlying conceptions of mind wandering suggest. Seli et al. (2018) offer an exhaustive description of a wide range of mind-wandering definitions.

Oftentimes, various definitions overlap. But imagine sitting in a lecture with the current task goal being understanding its topic, and yet willfully thinking about your upcoming birthday party. Following, for example, the unintentional-thought definition, you are not mind wandering because your birthday-party thoughts occur with intention. Following the TUT definition, you *are* mind wandering, because your birthday party is unrelated to the lecture's topic. It seems that the term mind wandering itself should rather be seen as an umbrella term subsuming similar, yet slightly differing constructs. The term was introduced by Smallwood and Schooler (2006) to, among other things, unify such constructs and elevate the status of the research topic, which it did quite successfully, as the research field "exploded" starting 2006 (Callard et al., 2013; Weinstein, 2018).

The variety of different labels already demonstrates that mind wandering is not a unitary construct, neither concerning the definition and theoretical understanding of the phenomenon, nor the resulting operationalization in empirical research (Seli et al., 2018). In order to reconcile this heterogeneous conceptual structure, Seli and colleagues suggested that "mind-wandering is best considered from a family-resemblances perspective; that is, as a heterogeneous, fuzzy-boundaried construct that coheres amid patterns of overlapping and nonoverlapping features" (p. 1). Acknowledging this heterogeneity and accepting that one single definition does not fully capture the variety of the construct, but yet in need of a specific conceptualization as a working basis, I will refer to mind wandering as TUT and will use both terms synonymously within this dissertation. Thereby, following the recommendations of Seli and colleagues, I specifically conceptualize the concept of mind wandering, making a classification within the literature possible and constraining my theoretical conclusions to this exact definition. I chose the TUT definition because it helpfully allows the experimental researcher to set up a certain (stable across participants, when desirable) task context and to therein investigate and identify negative as well as positive consequences of drifting thoughts, as is desirable within an experimental context. Seli and colleagues further argued that it allows investigating the wandering mind as an everyday phenomenon by recreating daily-life activities, such as reading or studying, giving practical relevance to findings obtained applying the TUT conceptualization. From a functional perspective, the TUT definition further is precise and concrete, facilitating the classification of reported thoughts in psychological studies.

### *1.2 Subjective In-The-Moment Mind-Wandering Assessment*

Human thought is externally not observable. We cannot see what others are thinking. Therefore, in order to capture the phenomenon of mind wandering, researchers often ask participants to self-report on their own internal states (Weinstein, 2018). The validity of self-reports is a heavily discussed topic in general psychological research. On the one hand, it was argued that people have only limited conscious access to their own cognitive processes and that self-reports underlie subjective distortions and cause reactive behaviors (Nisbett & Wilson, 1977). On the other hand, it was found that thought protocols strongly correlate with objective measures and that the self-reports' distorting influence is only weak (Ericsson & Simon, 1980).

In mind-wandering research, the self-report method can be broadly divided into two approaches (Weinstein, 2018) when it comes to measuring the phenomenon in the moment it is happening (i.e., *online*). Researchers employing the *self-caught* approach ask participants to monitor their own thoughts and to report, often via a simple button press, whenever they consider their thoughts as mind wandering. It is discussed whether the self-caught approach produces estimates of off-task thought *awareness* rather than *frequency*, because participants need to be aware of their current thought processes in order to observe and categorize drifting thoughts (Schooler, 2002; Schooler et al., 2004). Researchers employing the *probe-caught* approach, which is most widely applied in mind-wandering research, do not have to rely on meta-awareness to such an extent (Smallwood & Schooler, 2006). So-called *thought probes* allow for the collection of self-reports concerning participants' current thoughts at random or specific points in time during an empirical investigation. While working on a task, such probes interrupt participants and ask them to respond to short questions concerning their current thought experiences. Thus, thought probes are able to catch participants mind wandering sometimes even before they become aware of it, producing a more reliable estimate of off-task thought frequency (Schooler et al., 2011).

In recent years, the question of the validity of such thought probes emerged. At least three of their features (Ericsson & Simon, 1980; Schooler & Schreiber, 2004) speak for their eligibility as one of the main data sources within mind-wandering research. First, thought probes do not suffer from retrospective distortion, as they represent direct and immediate prompts to report on current thoughts. Second, on the part of the participants as well as the researchers, ambiguity is reduced when a discrete decision is required. Third and finally, no further knowledge, or expert knowledge, is required on the part of the participants since they merely have to report thought descriptions without speculating about the cause of their current mental state. And indeed, the validity of thought probes is supported by several studies showing that mind wandering reported via this method varies with situational influences, such as alcohol intake (Sayette et al., 2009), sleep deprivation (Poh et al., 2016), or hunger (Rummel & Nied, 2017). Further, mind wandering reported via thought probes correlates with behavioral and physiological markers, such as *reaction times* (Csifcsák & Mittner, 2017; McVay & Kane, 2012a; Seli, Cheyne, et al., 2013), *eye movements* (for a review on eye movements during mindless reading, please see Manuscript 1), and the activation of the *Default Mode Network* (Andrews-Hanna et al., 2014; Mittner et al., 2016), which subsumes brain regions that have been connected to resting as well as mind-wandering states. It is, thus, not surprising that the probe-caught approach emerged as the most popular thought measurement tool in recent years (Weinstein, 2018).

### *1.3 Deleterious Effects of Mind Wandering*

As mentioned above, the phenomenon of mind wandering has been given many names, which are not limited to specific scientific terminology, but also include informal, colloquial labels such as *mind lapses* or *cortical idling*. Together with terms such as *cognitive failure*, *rumination*, or *attentional failure,* a rather negative connotation is revealed, already hinting towards a detrimental and costly side of mind wandering. And indeed, TUTs are often associated with performance decrements within various domains. When minds drift, the risk of comprehension failure during reading increases (Feng et al., 2013; McVay & Kane, 2012b; Schad et al., 2012; Schooler et al., 2004; Unsworth & McMillan, 2013). In typical mind-wandering-during-reading experiments, participants are asked to read a given text while their thoughts are probed from time to time to assess mind-wandering levels. High levels are most often found to be related to worse performance on subsequent comprehension tests, affecting not only text information presented right before a thought probe (Reichle et al., 2010; Smilek et al., 2010), but going far beyond: Due to mind-wandering episodes, readers are found to experience trouble connecting single events to a full situational model of a story, leading to difficulties when, for example, identifying the villain in a crime novel (Smallwood, McSpadden, et al., 2008). Besides reading, one of the tasks employed most often to study detrimental effects of mind-wandering is the Sustained Attention to Response Task (SART). In this go/no-go task, participants are asked to respond to frequent non-target stimuli and to withhold responses to infrequent target stimuli. Errors on the SART have been connected to mind-wandering occurrences so frequently that they have even become a common behavioral marker of TUT episodes (Cheyne et al., 2009; Mrazek, Smallwood, & Schooler, 2012; Smallwood et al., 2004).

TUT occurrence is further associated with performance decrements in working memory tasks and intelligence tests (McVay & Kane, 2009; Mrazek, Smallwood, Franklin, et al., 2012), and, more generally, with disruptions to learning and lower academic achievement (Farley et al., 2013; Seli, Wammes, et al., 2016; Wammes & Smilek, 2017). Higher mind wandering levels assessed in the laboratory have even been connected to lower scores on the Scholastic Aptitude Test (SAT), that had been taken by participants one to three years before (Mrazek, Smallwood, Franklin, et al., 2012), highlighting drifting thoughts' negative influence within an educational setting. Of course, mind wandering not only happens when people are sitting at their desks, in front of a laboratory computer, or in a comfortable reading chair. Applied research has tracked mind-wandering deficits from the laboratory into daily life (McVay et al., 2009) and has identified TUTs as a potential safety hazard. When driving a car, wandering thoughts can quickly become dangerous (Albert et al., 2018; Lin et al., 2016; Yanko & Spalek, 2014), as, among other things, braking reaction times increase and headway distances decrease when thoughts are off-task.

However, not only performance suffers when the mind wanders. Off-task thoughts are furthermore associated with negative emotional consequences, as the title of a much cited research article (Killingsworth & Gilbert, 2010) already suggests: "A wandering mind is an unhappy mind." In this study, time-lagged analyses showed that mind-wandering episodes preceded negative mood in daily life, at least hinting towards such a causal relationship. Other studies suggest that a low mood causes people to engage in more self-referential mind wandering (Smallwood et al., 2009; Taruffi et al., 2017). It is thus possible that the relationship between mind wandering and mood is bidirectional. Concerning TUT content, mood decreasing effects are found particularly often when the mind wanders to the past, sometimes even when the thought content itself is positive (Ruby et al., 2013; Smallwood & O'Connor, 2011; Stawarczyk, Majerus, et al., 2013).

Solely concentrating on such far-reaching negative consequences of mind wandering, one can quickly get the impression that TUTs should be inhibited at all times. And indeed, intervention options exist, which aim at decreasing mind-wandering frequencies in daily life including mindfulness trainings based on meditation or mindful breathing (Bennike et al., 2017; Mrazek et al., 2012; Rahl et al., 2017). However, a different strategy, focusing on regulating rather than simply decreasing TUTs, becomes desirable or even necessary as soon as TUTs are found to not only have costs but also benefits, as the first should be minimized and the latter maximized.

## 2 Adaptive Mind Wandering

### *2.1 Beneficial Effects of Mind Wandering*

Since mind wandering is found to be both ubiquitous and deleterious, it probably comes as good news to all of us wandering companions that the phenomenon has also been related to beneficial effects (Mooneyham & Schooler, 2013; Smallwood & Andrews-Hanna, 2013). As early as in the 1970s, the facet of *positive constructive daydreaming* was introduced by Singer (1975). Challenging the one-sided impression that drifting thoughts come at nothing but high costs, he characterized this specific style of daydreaming as planful and creative. At a time when wandering thoughts were even associated with psychopathology (McMillan et al., 2013), he shaped a new, adaptive view of the phenomenon, and his ideas still influence current research in the field to this day.

From an adaptive perspective, it seems unlikely that people would spend up to 50 percent of their waking hours mind wandering (Killingsworth & Gilbert, 2010; Klinger, 1999), if drifting thoughts were not functional to at least some extent. And indeed, studies investigating the adaptive value of mind wandering support Singer's (1975) positive-constructive-daydreaming framework. Recent findings demonstrated that mind wandering can foster skills such as future and autobiographical planning (Baird et al., 2011; D'Argembeau et al., 2011; Smallwood et al., 2009), as well as social problem solving and socio-emotional adjustment (Poerio et al., 2016; Ruby et al., 2013). In addition, creative incubation and creative problem solving have been found to benefit from drifting thoughts (Baird et al., 2012; Gable et al., 2019; Leszczynski et al., 2017; Preiss et al., 2016; Zedelius & Schooler, 2015), supporting the widely-held belief that creative thinkers often let their minds wander (but see section 2.3.3). Other positive effects of mind wandering have been proposed, but still need to be examined. For example, Singer (1966) suggested that daydreaming can offer relief from boredom. Having to work on an uninteresting, tedious task, mind-wandering episodes offer a temporary way out. They might also function as mental breaks, during which people "recharge", as well as serve as a dishabituating activity by temporarily taking people away from a current task and enabling a "fresh look" when returning to it (Schooler et al., 2011). Furthermore, using mind-wandering episodes to mentally travel through time might allow us to bring in line our past and imagined future selves to the person we are today (Smallwood & Andrews-Hanna, 2013). Such processes might provide us with a sense of self-identity, meaning, and continuity across time (Prebble et al., 2013).

## *2.2 When to Wander, When Not to Wander?*

**2.2.1 Task-Difficulty as a Moderator of Mind-Wandering Frequency.** With more and more evidence emerging for negative as well as positive effects of drifting thoughts, both sides of the medal needed to be brought in line. To reconcile costs as well as benefits of mind wandering, Smallwood and Andrews-Hanna (2013) argued that the relationship between TUTs and their outcome is context dependent within their *context-regulation-hypothesis*. The authors stated that under attention-demanding conditions (e.g., when taking a difficult exam), mind wandering is unproductive and inefficient due to it representing a potential source of error for the task at hand. In contrast, low attentional task demands allow an individual to successfully perform a given task as well as to simultaneously engage in mind wandering. Such conditions enable drifting thoughts to come at no, or rather low and tolerable, costs and thus make room for beneficial effects of mind wandering to unfold. To make use of TUT benefits and to prevent costs, one should therefore know when to wander and when not to wander. And indeed, two patterns emerge from previous mind-wandering research, supporting the assumption of the context-regulation hypothesis (Smallwood & Andrews-Hanna, 2013; Smallwood & Schooler, 2015): First, there are high performance costs of mind wandering during demanding tasks. Second, TUTs are more prevalently found under conditions with low attentional demands. More direct evidence for the context-regulation hypothesis comes from Rummel and Boywitt (2014), who found that people with higher *working memory capacities* in particular are flexible in their adjustment of mind wandering to situational demands. People are able to coordinate on- and off-task thoughts in such a way that TUTs are more likely to occur under undemanding task conditions and less likely to occur during demanding tasks. Notably, this adjustment ability benefits task performance when task demands increase. Further support for the idea of a flexible TUT adjustment comes from C*ognitive Control Theories* concentrating on the allocation of cognitive control (Shepherd, 2019), such as the *Expected Value of Control Theory* (Shenhav et al., 2013; Shenhav et al., 2016). According to this theory, costs and benefits of directing cognitive control towards a current task are weighted against each other to decide how much control to exert. If costs and/or benefits of succeeding in the current task are low (e.g., during an easy task whose fulfillment is not personally relevant), only low levels of cognitive control are exerted. In turn, this could allow possibly adaptive mind wandering processes to "take over," particularly if the current mind-wandering content is personally relevant or promises higher benefits than current task fulfillment. In a very similar vein, Kurzban et al. (2013) argued that cognitive resources, which are not urgently required for current task-goal fulfillment, can and will be allocated to another task. Thinking about, or mentally working on, another task, can in turn be considered mind wandering, at least according to the TUT definition. Such a reallocation of cognitive resources and thoughts enables maximized combined task-goal fulfillment.

In our everyday life, it would come in handy if we were able to regulate our mind-wandering behavior in such a way. Given that thought adjustment could allow us to make use of mind-wandering benefits while keeping mind-wandering costs low, the importance of identifying cognitive (like working memory capacity, see Rummel & Boywitt, 2014) and other dispositional factors which facilitate and control adjustment processes becomes increasingly clear. However, this is not merely an individual differences approach, but rather a complex interplay of person- and situation-related factors. Taking the study by Rummel and Boywitt (2014) as an example, the level of task demands (situation-related) specifies the appropriateness and the costs of mind wandering. In turn, working memory capacity (person-related) contributes to the ability to regulate mind-wandering behavior in accordance to task demands. Thus, besides person-related trait-variables, situations allowing mind-wandering processes to take over, such as those with low task difficulty (Rummel & Boywitt, 2014), have to be identified. Factors, which describe such situations, might include but are not limited to the level of a task's personal relevance, the promise of expected financial or other incentives, or the level of interest in the given topic. In Manuscript 1, we built on the work by

Rummel and Boywitt (2014) and focused on *meta-cognitive beliefs*[1] of task difficulty as a potential moderator of mind-wandering frequency.

**2.2.2 Meta-Cognitive Beliefs as a Moderator of Mind-Wandering Frequency (Manuscript 1[2]).** *Meta-cognition* can be defined as the knowledge about one's own cognitive processes (Flavell, 1979). The interplay of *monitoring* and *control* functions of meta-cognition allows us to continually adjust ongoing behavior (Koriat & Goldsmith, 1996; Son & Schwartz, 2002). Thus, meta-cognitive beliefs about task demands might facilitate targeted employment of mind-wandering adjustment processes: Imagine a situation in which you are studying for a test. You usually have a "feeling" about the test's difficulty, which depends on the subject matter, your current abilities concerning the subject, and other factors. Expecting the upcoming test to be very demanding might lead to the direction of cognitive control towards the learning material (Shepherd, 2019), blocking out intruding TUTs. Using the terms of a highly influential framework of executive functions (Miyake & Friedman, 2012), one could argue that *inhibition* processes take over, making it possible to mainly aim attention at and perform well in the current learning task. In such a situation, mind-wandering costs would outweigh its benefits (ignoring further moderators such as personal interest or relevance), making a strategy to keep costs low decisive. Then again, knowing that an upcoming test will not be demanding might facilitate the employment of a different strategy. If attention, or cognitive control, could be directed towards task-related as well as task-unrelated thoughts, people could benefit from the positive effects of mind wandering, while still keeping its costs low. The low demands of the task make a *shifting* strategy possible (Miyake & Friedman, 2012) and thought processes centered on studying and, for example, planning tonight's dinner could alternate.

To examine the influence of meta-cognitive beliefs about task difficulty on mind wandering, we asked participants to read two book chapters for up to 30 minutes. We manipulated expectancies regarding an upcoming reading-comprehension test by telling one

---

[1] The main focus of Manuscript 1 was to validate eye-movement measures as indicators for mindless reading (see section 4.2). However, we also aimed to apply these indicators to examine meta-cognitive knowledge as a possible driving factor for mind-wandering adjustment. Including this research question within our validation study further allowed us to test for convergent validity between objective (eye movements) and subjective (self-reports) mind-wandering measures.

[2] Steindorf, L., & Rummel, J. (2020). Do your eyes give you away? A validation study of eye-movement measures used as indicators for mindless reading. *Behavior Research Methods, 52*(1), 162-176. https://doi.org/10.3758/s13428-019-01214-4

group of participants that the test would be very difficult while the other group was told that it would be fairly easy. To create credibility, we embedded the manipulation within a background story concerning the development of a standardized reading comprehension test for an adult population. During self-paced reading, participants were asked to report on their current thoughts via ten thought probes, which were allocated over the total reading time according to a specific pattern (section 4.2). A retrospective mind-wandering questionnaire, a manipulation check concerning difficulty expectations, as well as a moderately difficult reading-comprehension test followed for both groups. We assumed that participants in expectation of a difficult reading comprehension test would mind wander less during reading compared to participants with low-difficulty expectations. In addition, we expected less mind wandering to result in better test performance.

Meta-cognitive beliefs about task difficulty were successfully induced, as revealed by the significant differences in expected task difficulty reported during the manipulation check. However, these beliefs did not result in group differences concerning mind-wandering behavior. Both experimental groups reported the same amount of TUTs, online as well as retrospectively. However, within the group expecting a difficult test, we found difficulty expectations to negatively correlate with online and retrospective TUTs, while there was a near-zero correlation within the group expecting an easy test. Additionally, both groups performed equally well on the reading-comprehension test. Still, overall, more off-task thoughts went along with worse test performance, which is well in line with the literature on mind-wandering costs[3].

Our assumptions regarding the relationship between meta-cognitive beliefs about test difficulty and mind-wandering behavior were not supported by our data. The relationship between meta-cognitive beliefs and mind wandering might be more complicated than assumed. Concerning our study, we suggested in retrospect that extrinsic as well as intrinsic motivational factors might have been at play. In the experimental group expecting a difficult comprehension test, thoughts during reading might have been influenced by the high difficulty expectations acting as extrinsic motivation to perform well. This extrinsic motivation might have overshadowed intrinsic motivational factors, such as interest in the topic. As there was still variance concerning the test-difficulty expectations within this group (as the manipulation check revealed), we concluded that the higher these expectations, the

[3] For a more detailed description of the experiment's materials and results, please see Appendix A1.

more TUTs might have been inhibited so that they would not disturb reading comprehension (see our a priori assumptions). However, lacking such extrinsic motivation, low-difficulty-expectation participants' thoughts might have been more prone to being influenced by intrinsic motivational factors, possibly explaining the lack of a link between test-difficulty expectancies and mind wandering within this group. Factors such as topic interest might have superseded our experimental manipulations in the direction that highly interested participants might have nevertheless inhibited TUTs, even though there was no urgent need to do so due to beliefs about test difficulty.

Nonetheless, this reasoning remains a post hoc explanation of our data. The existence and interplay of various situation- and person-related factors influencing mind-wandering adjustment might hamper the isolated examination of one single factor. Identifying main effects and interactions of such factors is still a work in progress. Situations in which mind wandering is appropriate and even desirable need yet to be identified as well as trait factors that control or benefit adjustment processes in accordance with current situational demands. Further insights could expand the context regulation hypothesis (Smallwood & Andrews-Hanna, 2013) in the direction of a more comprehensive framework of mind-wandering adjustment. Experiments such as ours should be extended and potential further influence factors besides the one(s) being manipulated should at least be assessed and statistically considered. In a follow-up experiment, one could, for example, assess personal interest in the text being read and include this measure as a covariate.

### *2.3 In Search of Further Beneficial Effects of Mind Wandering*

**2.3.1 Memory Benefits Due to Mind Wandering (Manuscript 2[4]).** Although, in Manuscript 1, we were not able to identify situations with metacognitive low-difficulty *beliefs* as occasions for increased mind-wandering frequencies, situations with actual low task difficulty do indeed allow the mind to wander (section 2.2.1). However, just because one *can* and *does* mind wander during easy tasks, does not mean that anything good has to come of such mind-wandering episodes. Thus, besides identifying situations in which mind wandering occurs most frequently and with no or only low costs, it should be of great interest to examine positive consequences of drifting thoughts. As mentioned above (section 2.1), researchers have demonstrated positive effects of TUTs in various domains such as *creativity* (e.g., Baird

[4] Steindorf, L., & Rummel, J. (2017). "I should not forget the apples!" – Mind-wandering episodes used as opportunities for rehearsal in an interrupted recall paradigm. *Applied Cognitive Psychology, 31*(4), 424-430. https://doi.org/10.1002/acp.3328

et al., 2012; Gable et al., 2019), *future planning* (e.g., Baird et al., 2011; Smallwood et al., 2009), and *social problem solving* (Ruby et al., 2013). *Memory* represents yet another domain that may benefit from mind-wandering processes. The superiority of spaced versus massed learning has been the subject of a large body of psychological studies (e.g., Cepeda et al., 2006; Donovan & Radosevich, 1999; Godbole et al., 2014). As Ebbinghaus (1964, p. 89) already stated more than fifty years ago, "a suitable distribution of [repetitions] over a space of time is decidedly more advantageous than the massing of them at a single time". In Manuscript 2, we suggested that mind-wandering episodes might offer an opportunity for such spaced learning. People could be able to adaptively make use of their wandering thoughts by including material-rehearsal episodes within a different ongoing task. Memorizing materials, such as a grocery-shopping list or the date and time of an upcoming appointment, is part of our everyday life. Using mind-wandering episodes for this kind of rehearsal could constitute an efficient way of improving memory performance and would be well in line with the cognitive-resources-allocation ideas introduced in section 2.2.1.

When, for example, reading a book chapter, understanding its content can be defined as the present task goal. However, this goal might not be the only active task goal. As Klinger (1967, 2013) suggested, people usually have more than one *current concern*, meaning that there are goals extending beyond the here and now. For example, keeping to-be-studied materials, such as a grocery-shopping list, in mind could be considered a further task goal, which might play a subordinate role in the moment, but will become important in the future. In a situation with active present and future task goals, one can allocate thoughts to either the present (i.e., on-task thought) or the future task goal, or to completely unrelated issues. Therefore, in Manuscript 2, we defined *goal-related mind wandering* as TUTs, which are used for the maintenance of future task goals. Conversely, *goal-unrelated mind wandering* can be characterized as TUTs, which do not concern any present or future task goals, such as those concerning the weather, or the reminiscence of a past vacation. Considering the context-regulation hypothesis (see above and Smallwood & Andrews-Hanna, 2013), people should benefit from engaging in goal-related mind wandering when the situation requires and/or allows it. More precisely, during undemanding tasks, people might easily be able to allocate thoughts towards a pending task goal, such as the rehearsal of a shopping list. The situation *allows* goal-related mind wandering because the ongoing task does not require the entirety of attentional resources. The situation also *requires* goal-related mind wandering, because the pending task goal will become important in the future. Consequently, people should in fact allocate a certain amount of thoughts towards the future task goal to gain the memory-benefits of rehearsal. Furthermore, it seems advisable to prevent general, goal-unrelated mind wandering in order to not sacrifice the active task goals. First evidence for our assumptions concerning thought allocation comes from a study by Rummel et al. (2017), who revealed that people show lower levels of goal-unrelated mind wandering when a special kind of future task goal (i.e., a pending prospective-memory intention) is present.

Accordingly, in our work, we aimed to investigate whether people use mind wandering in an adaptive way when, besides a present task goal, a future memory-task goal is present. Further, we intended to examine the nature of rehearsal strategies. Put more precisely, we took a closer look at the persistence of goal-related mind wandering over time. Finally, to demonstrate the actual success of rehearsal during mind wandering, we examined people's memory for the to-be-remembered materials, which can be described as the degree of future task-goal fulfillment. In our experiment, we asked the participants to study a grocery-shopping list and to recall half of the items after the study phase. The recall phase was framed as a "trip to the supermarket". Intending to activate a future task goal in one experimental condition (*interrupted condition*), we told participants that they would have to recall the other half later ("in another supermarket"). First, they would work on a different task, namely a moderately demanding letter-task. In order to not activate this future task-goal in the other experimental condition (*finished condition*), participants were told that the recall task was completed before they started working on the letter task. While participants worked on the letter task, we assessed the amount of mind wandering as well as participants' thought contents in order to test for effects of the future-task goal (in)activation. To this end, we employed ten thought probes, which interrupted the task, to ask participants to categorize their current thoughts as being either on- or off-task and to type in a brief description of their current thoughts' content. Finally, in order to measure the degree of future task-goal fulfillment, we asked all participants to recall the second half of the shopping list (see Figure 2 for a depiction of the procedure).

The results showed that all participants performed equally well during the first recall phase, which was not surprising given that the experimental manipulation took place only after this phase. Working on the following letter task, all participants performed similarly well and showed similar levels of overall self-categorized mind wandering. Still, participants in the interrupted condition showed higher levels of goal-related mind wandering towards the shopping list than participants in the finished condition, as assessed by the short thought descriptions[5]. That is, the activation of a future task goal influenced the proportion of goal-related and goal-unrelated mind wandering without influencing the overall level of TUTs. Furthermore, the likelihood for the occurrence of goal-related mind wandering generally decreased over time. However, this time course differed between conditions: Participants in the interrupted condition showed a stronger and more persistent mental engagement with the future task goal than participants in the finished condition, who stopped thinking about the shopping list early on. Not only did interrupted participants show a more efficient rehearsal strategy, they also performed better in the second recall phase[6].

We interpreted our results in such a way that the increased and more persistent mental engagement in goal-related mind wandering may indeed have been a successful rehearsal strategy for the fulfillment of a future task goal. Using mind-wandering episodes to rehearse a shopping list might actually prevent people from forgetting important items at the supermarket. Interestingly, our interrupted participants displayed a highly efficient thought-adjustment behavior. They did not sacrifice any of their on-task thoughts (i.e., present-task-goal-related thoughts) for the maintenance of the future task goal. Instead, they showed lower levels of general, goal-unrelated mind wandering. This flexible thought adjustment might have been the reason why interrupted participants' performance on the current task did not suffer in the presence of a future task goal. People seem to be able to adjust their mind-wandering content according to active future task goals while still fulfilling their present task goals. They are able to efficiently allocate thoughts to present as well as future task goals. Thus, in Manuscript 2, we demonstrated mind-wandering benefits in the domains of memory and goal maintenance.

**2.3.2 No Problem-Solving Benefits Due to Mind Wandering (Manuscript 3[7]).** Aside from the maintenance of memory content, finding a solution to a pending problem represents another key future task goal in daily life. In Manuscript 3, we examined whether people would allocate thoughts towards a pending decision problem during an incubation interval and whether this mental occupation with the problem would foster a solution. Put more precisely, we took a closer look at the so-called *Unconscious Thought Effect*

---

[5] Participants' brief thought descriptions were rated as to whether they related to the shopping scenario or not.
[6] For a more detailed description of the experiment's materials and results, please see Appendix A2.
[7] Steindorf, L., Boywitt, C. D., & Rummel, J. (2020). *A fresh look at the Unconscious Thought Effect: Using mind-wandering measures to reveal thought processes in decision problems with high information load.* Manuscript submitted for publication.

(Dijksterhuis et al., 2006; Dijksterhuis & Nordgren, 2006), which manifests itself in the form of better solutions to complex problems supposedly due to previous unconscious thought in contrast to conscious problem-elaboration. When faced with the decision which of four different apartments to rent or cars to buy, one has to integrate a multitude of attributes. According to the *Unconscious Thought Theory*, the high information-processing capacity of the unconscious system should be better suited to make such a decision than the conscious system, which suffers from a low information-processing capacity. In a standard unconscious thought paradigm, participants are introduced to several objects and corresponding attributes (e.g., "Apartment 1 has a balcony.", or "Apartment 3 does not have a washing machine."). A higher number of positive attributes characterizes objectively better objects. After being presented with and before being asked to choose from the list of objects, participants work on a distraction task or engage in conscious thought concerning the decision. The distraction task, especially when it is highly demanding, is supposed to only allow for unconscious occupation with the decision problem, as it otherwise engages the participants' conscious resources. Evidence for an Unconscious Thought Effect comes in form of better decisions or more accurate object evaluations after distraction periods in contrast to conscious-elaboration periods without a distraction task. However, it should be noted that the Unconscious Thought Theory has been heavily criticized and, despite many successful replications, the effect does not always replicate (Acker, 2008; Calvillo & Penaloza, 2009; Newell & Rakow, 2011; Rey et al., 2009).

In Manuscript 3, our aim was to use mind-wandering measures to reveal thought processes in a standard unconscious thought paradigm, as one of the still open questions concerning the Unconscious Though Effect is the one addressing cognitive processes underlying the effect (Strick et al., 2011). In all three experiments of Manuscript 3, thought probes and/or retrospective mind-wandering questionnaires were employed to measure thought-contents. Besides offering insights into the cognitive processes during distraction and conscious-deliberation periods, measuring participants' thoughts was supposed to allow us to test for an alternative explanation for Unconscious Thought Effects in standard paradigms: The presentation of object-attribute combinations might activate a goal to solve the current decision problem. When the solving process is interrupted by a distraction task in the respective experimental conditions, the goal might remain pending and active (Watkins, 2008), which might further result in an increased mental occupation with the problem (Zeigarnik, 1927). Such mental occupation can be considered mind wandering and might, in turn, foster problem solving (Baars, 2010). Wandering apartment thoughts, aware or unaware[8], might be what is referred to as “unconscious” within the Unconscious Thought Theory. More precisely, people could integrate short problem-solving episodes into the distraction task. Furthermore, such adaptive mind-wandering processes should be better able to unfold during low-demanding distraction tasks, as high-demanding tasks might act as a competitor for attentional resources, suppressing adaptive TUTs[9], thus contradicting some of the Unconscious Thought Theory’s assumptions.

In three experiments, participants worked on a complex apartment-evaluation task, as exemplarily described further up in this section. We separated the evaluation task’s presentation and evaluation phases by periods of different activities varying between participants, including demanding as well as undemanding distraction tasks and conscious apartment elaboration[10] (see Figure 2 for a depiction of the procedure). Conscious-thought conditions were supposed to test for the existence of an Unconscious Thought Effect by representing a criterion for comparison to distraction-task conditions. Distraction-task conditions were further supposed to test for our alternative explanation: We expected more general and more apartment-related TUTs for participants working on an undemanding distraction task compared to participants working on a demanding distraction task (e.g., Rummel & Boywitt, 2014). In turn, participants working on the undemanding task should perform better on the later evaluation task due to previous adaptive mind-wandering processes. Assessing thoughts during (via thought probes) and after (via retrospective questionnaires) the distraction period was supposed to allow us to relate the respective thought frequencies to evaluation performances.

An Unconscious Thought Effect, that is better evaluation-task performance for distraction-task groups than conscious-thought groups, was present in the first two experiments, but not in the third. Further, lower distraction-task demands resulted, in general, in more mind wandering towards the apartment-evaluation task as well as towards other distraction-task unrelated issues than higher distraction-task demands. However, higher

---

[8] Please see Appendix A3 for a more extensive discussion of awareness issues.

[9] The assumption that the highest amount of apartment thoughts is probably present in conscious-thought conditions does not imperatively challenge our alternative explanation for the Unconscious Thought Effect for two reasons: First, too much problem deliberation could have destructive effects, as the problem might be “thought to pieces” (Waroquier et al., 2010). Second, conscious-thought conditions might not benefit from a potential fresh look at the pending problem due to a dishabituating function of the distraction task.

[10] For a full description of our methods, experimental conditions, and the underlying theoretical assumptions and hypotheses, please see Appendix A3.

apartment-thought rates did not go along with a better performance in the evaluation task. On the contrary, less apartment-related mind wandering seemed to be associated with better evaluation performance.[11] However, in two experiments, this relationship was confounded by thought-probing effects, which I will refer to in section 3[12].

Even though we employed similar methods in all three experiments, evidence concerning the Unconscious Thought Effect was mixed, adding to the discrepancies in the currently ongoing discussion concerning the size of the effect. Additionally, we confirmed that high distraction-task demands compete with (adaptive) mind-wandering processes for attentional resources. However, there was no evidence for a connection between the level of apartment-related mind wandering and the evaluation performance. Thus, in general, there was no evidence for mind wandering as an alternative explanation for the Unconscious Thought Effect and no evidence for problem-solving benefits due to TUTs.

**2.3.3 No Creativity Benefits Due to Mind Wandering (Manuscript 4[13]).** It is not always a precise problem with one single correct solution, as the one examined in Manuscript 3, that occupies our minds. Sometimes we have to come up with novel and useful ideas within a wide-open space of possible solutions. It is common belief that creative ideas arise when minds are allowed to wander freely (off-task thought) and that focusing too intently on a problem (on-task thought) does not foster a creative solution. Referring to this belief, recent research has discussed the beneficial effects of mind wandering for creativity and creative incubation. Manuscript 4 adds to this ongoing discussion.

According to the *explicit-implicit interaction model* of creative thinking (Hélie & Sun, 2010), creative incubation involves unconscious and implicit associative processes, which, while not exclusive to the wandering mind, might increasingly emerge during TUT episodes. And indeed, a positive relationship between mind wandering and creative performance was reported, for example, for professional writers and physicists, who indicated that many of their most significant ideas arose during mind-wandering periods (Gable et al., 2019). Also, increased mind-wandering tendencies were found to be connected to the occurrence of aha-experiences (Zedelius & Schooler, 2015), such as the popular Archimedes' "Eureka!"

---

[11] For a more detailed description of the experiment's results, please see Appendix A3.
[12] This summary of results is oversimplified as it ignores a further experimental manipulation, namely the existence or absence of thought probes during the distraction task. I will refer to this issue in section 3.1 by focusing on thought-awareness-inducing effects of such probes.
[13] Steindorf, L., Hammerton, H. A., & Rummel, J. (2019). *Mind wandering outside the box – About the role of off-task thoughts and their assessment during creative incubation.* Manuscript submitted for publication.

moment, as well as better performance on divergent and convergent thinking tests (Leszczynski et al., 2017; Preiss et al., 2016). Besides such correlational findings, a study by Baird et al. (2012) reported causal evidence for mind wandering benefitting creativity within an incubation paradigm. Engaging in an undemanding task during an incubation interval resulted in performance improvements on previously worked-on creativity tasks. Such an improvement was not present after engagement in a demanding task, a period of rest, or after having no break at all. Importantly, in a retrospective questionnaire, participants who had worked on the undemanding incubation task reported increased mind-wandering levels. The authors thus concluded that incubation mind-wandering fosters creativity. However, in a methodologically very similar study, Smeekens and Kane (2016) were not able to replicate this finding, as mind wandering during incubation did not predict post-incubation creativity. One major aim of Manuscript 4 was to resolve this inconsistency by focusing on mind-wandering assessment issues, which will be covered in section 3.1. At this point, I will focus on possible beneficial effects of mind-wandering within the domain of creativity.

In our study, participants worked on verbal as well as figural creativity tests (i.e., first creativity assessment) before being assigned to one of three incubation conditions[14] or a no-incubation condition. During incubation, all incubation participants worked on an easy letter task, which had previously been found to allow for fairly high mind-wandering levels (e.g., Baird et al., 2012). A retrospective assessment of incubation mind-wandering towards the creativity tests as well as towards other letter-task unrelated issues followed, before participants worked on both creativity tests once more (i.e., second creativity assessment). For participants in the no-incubation condition, this second creativity assessment directly followed the first assessment. Afterwards no-incubation participants worked on the letter task and filled out the retrospective mind-wandering questionnaire (see Figure 2 for a depiction of the procedure). In general, we expected creative incubation benefits, namely better performance in the second creativity assessment for incubation participants than for no-incubation participants. Further, and more specifically, we expected levels of incubation mind-wandering to differ between the three incubation conditions (see section 3.1), with at least one of the incubation conditions showing higher TUT levels as well as better performance in the second creativity assessment than the baseline no-incubation condition to support the idea of mind wandering benefitting creative incubation. We differentiated between general TUTs (i.e.,

[14] The three incubation conditions were designed to test for effects of different mind wandering assessment methods and are not of interest here, but in section 3.1.1.

implicit associative processes) and creative TUTs (i.e., target-oriented and directly focused on creative solutions) to be able to test for influences of both kinds of mind wandering.

We did not find post-incubation creative performance benefits for any of the three incubation conditions when comparing their performance on the second creativity assessment (verbal as well as figural) to the no-incubation condition. Furthermore, neither via incubation-condition comparisons nor via correlational analyses did we find evidence for incubation mind-wandering being related to post-incubation creative performance[15]. Thus, our experiment supports the findings by Smeekens and Kane (2016) and we concluded that it might not be possible to increase creativity with mind-wandering manipulations. The mere opportunity to mind wander in an artificial lab environment might not be sufficient for creative incubation. However, we did not rule out the possibility that incubation mind-wandering might be found to be beneficial within paradigms producing reliable incubation effects and that mind wandering and creativity could be related on a trait level, as previous questionnaire research suggests (see above within this section).

**2.3.4 Discussion of the Search for Further Beneficial Effects of Mind Wandering [16].** Manuscript 2 introduced memory and goal maintenance as further domains which benefit from mind-wandering processes, adding to the current literature on adaptive mind wandering. However, in Manuscripts 3 and 4 we did not demonstrate beneficial effects of mind wandering within the domains of problem solving and creativity. As described in section 2.1, benefits of mind wandering have received increased attention in the past decade. However, as there are also studies that do not detect such positive effects, specific patterns and regularities emerging in the literature should be of great interest, beyond a pure listing of effects. Why and when do positive effects show? Is there *good mind wandering* and *bad mind wandering*? And more specifically, concerning this dissertation, why did we find mind-wandering benefits in Manuscript 2 but not in Manuscripts 3 and 4, although the studies' designs were quite similar (see Figure 2)? Is there a central difference between memory tasks and creativity or decision task, which could explain that we found positive effects of TUTs for one, but not the other? Trying to answer these questions, I will take a look at adaptive mind wandering from a

---

[15] For a more detailed description of the experiment's materials and results, please see Appendix A4.

[16] Individual discussions of the respective manuscripts' results can be found in the respective appendices. The current section aims at a rather collective discussion of patterns emerging from all three manuscripts (2-4).

content and a methodological perspective and I will further propose an individual-differences approach based on the assumption that not everyone mind wanders in the same way.

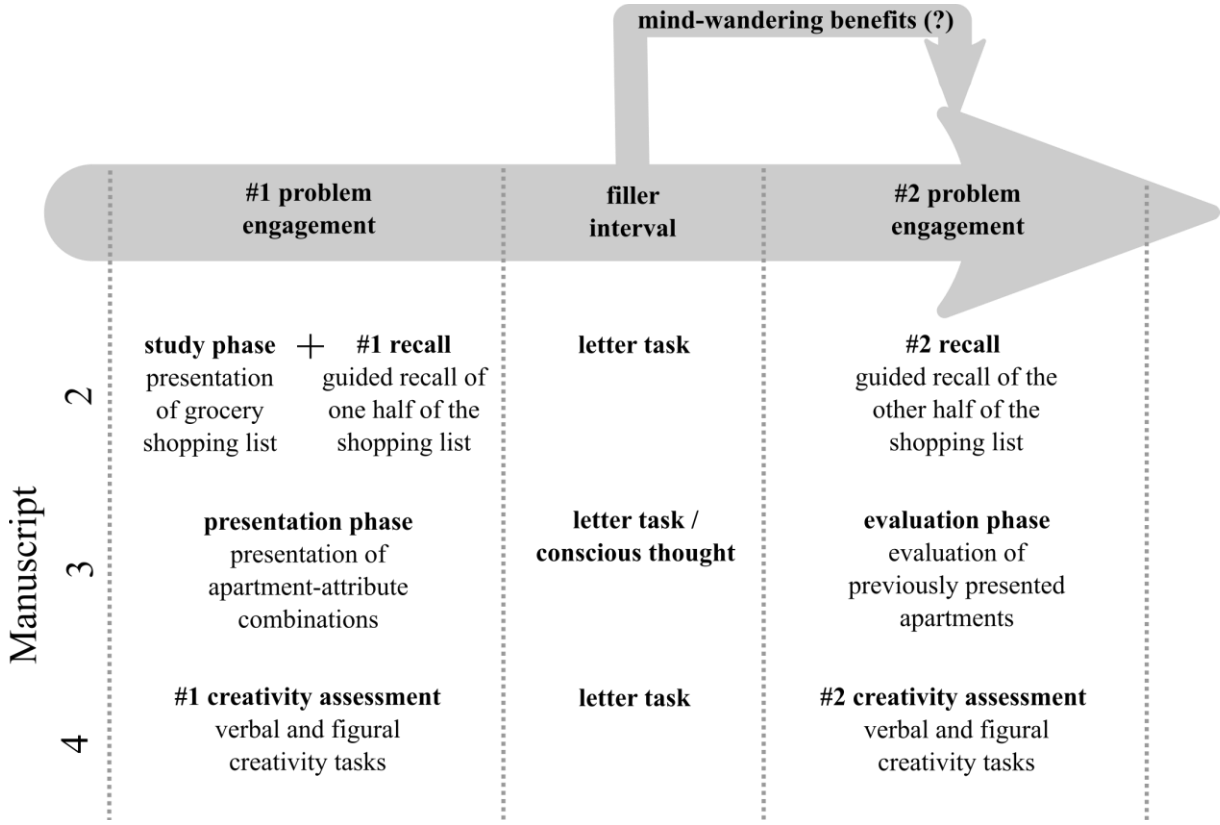


*Figure 2.* Simplified depiction of the study designs of Manuscripts 2, 3, and 4. In all studies there was a first problem engagement phase before participants worked on a filler task within the filler interval. We were interested in participants' mind-wandering behavior during this interval and expected it to be positively related to their performance in the following second problem engagement phase.

***2.3.4.1 Adaptive Mind Wandering From a Content Perspective.*** The specific topic of a mind-wandering episode determines its outcome valence to a great extent (*content regulation hypothesis*, Smallwood & Schooler, 2015). It sounds self-evident that positive mind wandering (pleasant, cheerful, planful, etc.) leads to positive effects. However, the picture is more diverse. Negative outcomes are often found when the mind wanders to the past, sometimes even when the thought content itself is positive (Ruby et al., 2013; Smallwood & O'Connor, 2011; Stawarczyk, Majerus, et al., 2013). In contrast, future thinking, which is very common during mind-wandering episodes (e.g., Smallwood et al., 2009; Stawarczyk et al., 2011), is found to go along with rather positive outcomes such as the improvement of mood, or future and autobiographical planning benefits (Baird et al., 2011; D'Argembeau et al., 2011; Ruby et al., 2013; Smallwood et al., 2009). Besides the temporal focus, personal interest in a current mind-wandering topic determines the valence of TUT outcomes. For example, drifting thoughts focused on topics, which people find highly engaging, are associated with an increase in positive mood (Franklin, Mrazek, et al., 2013).

Most of the aforementioned effects of specific TUT contents refer to mood. In addition to *temporal focus* and *interest,* and referring to positive effects on specific active tasks, further content-related factors such as personal relevance or goal-orientation could play an important role in determining the functional outcome of mind-wandering episodes. As a possible reason for the differences in results between Manuscripts 2-4, I propose a post-hoc explanation utilizing the interplay of such factors. Future orientation and personal relevance have already been found to go together (Stawarczyk, Cassol, et al., 2013): Future-related mind wandering is more personally relevant than non-prospective mind wandering. Also, most of our daily-life present and future goals are personally relevant to us. My reasoning within the scope of this dissertation is that *personally relevant and goal-oriented prospective TUTs* are beneficial when it comes to fulfilling a task goal: A task such as remembering a shopping list represents a potential personally relevant, future oriented goal. The shopping task in Manuscript 2 was operationalized in a way that resembles everyday grocery shopping. Everyone knows the feeling of being annoyed when cooking and realizing that one forgot to buy one of the most important ingredients. A failure in remembering parts of a shopping list has obvious future consequences in everyday life. A failure in coming up with creative solutions to verbal or figural creativity tasks (Manuscript 4) cannot as easily be translated into daily-life consequences, and such a task is probably not personally relevant to most people. Although many people might know the situation of having to choose between apartments or other objects, the way such a decision-task is operationalized in the unconscious thought paradigm (Manuscript 3) is rather artificial and might not evoke a feeling of personal (future) relevance or the goal to do well in the apartment evaluation within participants. Thus, shopping-task related TUTs in Manuscript 2 might represent the only mind-wandering content which can easily be applied to everyday life, has personal future relevance, and is strongly goal-related in comparison with creative thoughts (Manuscript 4) and apartment thoughts (Manuscript 3). However, the relation between such mind-wandering content features and positive mind-wandering outcomes remains a post-hoc explanation for the data and should be further tested by specifically manipulating the respective features. For example, a random word list without a goal-related (shopping) context might not be as continuously rehearsed within mind-wandering episodes as was the case in Manuscript 2. Further, tasks, which *naturally* lack personal relevance (as I propose is the case within Manuscript 3 and 4), could be incentivized to add an external relevance-factor.

Besides the specific thought content, the task formulation and the type of problem might also play a role. Referring to control theories, Shepherd (2019) argues that the mind might look for another active goal when the current task goal does not require all resources or is deemed insufficiently rewarding. This goal-search as well as the following attempt at goal-fulfillment might be easier and more successful if there is a very concrete active goal with a clear solution. The goal of memory or shopping-item maintenance (Manuscript 2) is a very specific one. The desired final state (recall of all shopping items) as well as the means to this end (rehearsal of the items) are known. These features might make goal switching fairly easy in comparison to rather vaguely defined final states such as a coming up with creative solutions (Manuscript 4). Moreover, even if switching to a creative goal was successful, solution approaches are even more vague. There is no known recipe for creativity as there is a recipe for memory maintenance (e.g., rehearsal). You cannot simply evoke or force creative thoughts, but you can make yourself, or a participant in the lab, rehearse items. The desired final state in the apartment task (Manuscript 3) may be clear to most participants (finding the best apartment). However, the solution approach is not that simple in a standard unconscious-thought paradigm. It might be considered a mixture of object-attribute memory maintenance, single-attribute evaluation, and the comparison of objects (including, at best, all object-attribute pairs) against each other. Such a solution approach is not as straight forward as simple maintenance of single grocery shopping items. *Concreteness* and *simplicity* of task formulation and goals, as well as of solution approaches, represent further content-related factors, which differ between Manuscripts 2-4 and might explain why there were mind-wandering benefits in one but not the others. These factors should be especially relevant for laboratory experiments because they only offer a limited time frame for the search of active goals and their fulfillment and thus for beneficial mind wandering (see also section 2.3.4.2).

***2.3.4.2 Adaptive Mind Wandering From a Methodological Perspective.*** Shortly summarizing the previous sections, TUTs might be beneficial for some tasks, but not for others. But how can we make sure to empirically capture benefits, given they exist? From a methodological perspective, one could argue that identifying benefits of drifting thoughts might require different approaches than identifying costs. When examining costs, the research focus lies mostly on TUT frequency: More TUTs lead to increased costs. Thought content is oftentimes ignored. However, when examining benefits, content might become a crucial component. Manuscript 2 built on this idea as it was a content-focused examination of beneficial mind-wandering effects. Focusing only on mind-wandering frequency would not have been insightful in the reported experiment[17]. Instead, short descriptions of the participants' thought contents revealed the adaptive nature of drifting thoughts used as opportunities for rehearsal. When examining costs, it is not as important whether people think about penguin knees or their last breakfast. It is the sole fact that people do mind wander, mostly independent of thought content, that influences current task performance. Thinking about penguin knees, however, does not help when it comes to planning the next trip to the supermarket and cannot be equated with shopping-list maintenance. Thus, more exhaustive methods assessing participants' thought content might be necessary when examining mind-wandering benefits.

A further difference between the examination of costs and benefits might be a temporal one. Costs often reveal themselves right away in the form of performance decrements. In contrast, benefits might rather tend to occur time-delayed or manifest themselves on a trait-based level only. Recent research connecting mind wandering with creative benefits illustrates the methodological difficulties that can occur when examining positive effects of mind wandering: The e*xperimental method* produced mixed effects concerning the causal relation between mind wandering during incubation intervals and later creative performance. In several experiments, Smeekens and Kane (2016) as well as our own research group (Manuscript 4) did not detect beneficial effects of mind wandering as were found by Baird et al. (2012). Manipulating state mind wandering in an artificial lab environment in order to examine positive effects of drifting thoughts thus seems to be a rather difficult endeavor, at least in the domain of creativity (but see Manuscript 2). We further do not know if creative benefits would have manifested themselves at a later time, as a laboratory experiment is bound to a very limited time period.

Correlational *content analyses* as well as *trait-based questionnaire studies*, however, support the idea of mind wandering fostering creative problem solving (e.g., Hao et al., 2014; Preiss et al., 2016; Zedelius & Schooler, 2015). To name one recent example, Gable and colleagues (2019) asked professional writers and physicists about their most creative idea of the day and the circumstances of its emergence. Thus, their participants' creative solutions were not time limited and not limited to state-creativity. The authors found that one fifth of such ideas emerged during mind-wandering episodes and that these ideas were more likely to

---

[17] In Manuscript 2, overall mind-wandering frequencies did not differ between experimental conditions and could thus not be associated with memory-performance benefits.

be experienced as “aha” moments compared to ideas created during on-task-thought episodes. This pattern within the domain of creativity (namely the existence of correlational effects in trait-based studies, but an unclear picture when it comes to experimental, state-based examinations) suggests that researchers investigating beneficial effects of mind wandering should be encouraged to take advantage of the methodological diversity available in the field, which goes far beyond simple questionnaire studies. For example, mobile-phone applications that transfer the experience sampling method to daily life (e.g., McVay et al., 2009) allow for the examination of time courses and might hint to causal relationships.

Ultimately, researchers should have in mind that benefits of mind wandering might reveal themselves under different circumstances than costs might do and should select their methodological approach carefully. Besides study-design issues and independent of the benefits-or-costs question, the methods and materials used to assess mind wandering online (e.g., thought probes) should also be added to this methodological perspective. All three manuscripts searching for mind-wandering benefits employed thought-probing methods, which might be considered a limitation as thought probes interrupt people’s trains of thought and make thought contents aware. This issue will be considered in more detail in section 3.

***2.3.4.3 Adaptive Mind Wandering Within an Individual-Differences Approach.*** Besides the question whether positive effects of mind wandering exist, the question of whether they exist for *everyone* should be considered. People differ in many traits and behaviors. It would thus be foolish to assume that all people mind wander in the same way and that mind-wandering benefits arise in the same manner for everyone. When it comes to mind-wandering benefits, employing an individual-differences approach can quickly become complex. Taking the creativity topic from Manuscript 4 as an example, only a specific *mind-wandering style* might improve creative performance, only *highly creative* people might benefit from mind-wandering processes, or only a specific *belief about the relationship* between mind wandering and creativity might lead to mind-wandering benefits. Individual differences might thus play a role at various levels. At the *mind-wandering level*, one could, for example, argue that people with higher manifestations of trait-level deliberate mind wandering (Seli et al., 2015; Seli, Risko, et al., 2016) or positive constructive daydreaming (Huba et al., 1981; Singer, 1974; Singer & Antrobus, 1970) practice a more constructive form of mind wandering and might be more prone to positive effects. However, at this point in time, not much is known about the characteristics of people who score high on questionnaires measuring such traits (but see Seli et al., 2015 for a relation to mindfulness facets and Zhiyan

& Singer, 1997 for a relation to personality variables) and, to my knowledge, no studies examining mind-wandering benefits have considered differences in mind-wandering traits.

At the *creativity level*, one could further argue that only highly creative people know how to adaptively make use of mind-wandering periods and might even have developed specific thought strategies which help them to come up with creative ideas. Thus, mind-wandering benefits might only be present for people scoring high on creativity measures. At the *intermediate level between mind wandering and creativity*, one could argue that only people who believe in the existence of a relationship between the two concepts show creativity benefits due to mind wandering. Such beliefs, or implicit theories, can be considered schemas, which simplify people's grasps of the external world and even guide cognition or behavior (Dweck et al., 1995; Fiske & Taylor, 1991; Kong & Jolly, 2019). Zedelius and Schooler (2017) suggested that some people might believe mind wandering to be valuable when it comes to creative insights and that these people might thus make better use of TUTs and experience beneficial TUT effects.

The ideas concerning individual differences discussed within the scope of the creativity example can be translated to other domains and present starting points for future research. Maybe only people with high problem-solving abilities know how to make use of mind-wandering episodes to solve previously presented problems. Maybe people who often mind wander intentionally can remember even more grocery shopping items than people who do not consider themselves frequent wanderers. As individual differences have been mostly ignored in the investigation of adaptive mind wandering, I believe this individual-differences perspective to be a very fruitful approach for future research.

## 3 Deleterious Effects of Thought Probes

All four manuscripts discussed in this dissertation employ the thought-probe method (section 1.2), mirroring its popularity in the mind-wandering field. As more and more researchers are currently using this approach to measure TUTs in the exact moment they are happening (Weinstein, 2018), the measurement method itself has become a research topic of interest. Although thought probes seem to be reasonably valid mind-wandering assessment tools (section 1.2), methodological problems shine through, starting with the observation that probes come in many shapes and forms (Weinstein, 2018). Studies using thought probes differ in features such as the wording and framing of instructions as well as of probes and response options, the number of response options (if any), or the time between probes. It has been argued that variations in such probe characteristics can lead to considerable differences in results (Robison et al., 2019): Honesty primes have been found to reduce self-reported mind wandering (Vinski & Watter, 2012). Furthermore, the probe rate seems to affect mind-wandering reports in the direction that people report higher TUT levels when the time between probes increases (Seli, Carriere, et al., 2013). Additionally, the framing of thought probes has been found to influence reported mind-wandering rates. Asking participants whether they were on-task can lead to fewer reported TUTs than asking whether they were off-task (Weinstein et al., 2018)[18].

***3.1 Thought Probes' Influence on Mind-Wandering Processes***

Besides effects of special thought-probe characteristics, effects of the mere presence of thought probes have become a topic of interest. Wiemers and Redick (2019) raised the question of whether people perform differently on a task when thought probes are employed. The authors found that such probes do not interfere with performance on an ongoing cognitive task. Furthermore, Robison et al. (2019) used a very similar task and found that performance did not differ between two experimental conditions varying in thought-probe frequency. Therefore, thought-probe presence does not seem to interfere with current task performance.

However, thought-probe presence might still alter thought or mind-wandering experiences during ongoing tasks. By continually disrupting the task, probes also interrupt the "train of thought", may it currently be on-task or off-task thought. Furthermore, going beyond purely interrupting effects, (repeatedly) asking people about their current mental states could increase their meta-awareness about thought contents and mind wandering behavior in general (Zedelius et al., 2015).

**3.1.1 Development of a New Paradigm Dissociating Between Interruption and Awareness Effects of Thought Probes (Manuscripts 3 and 4).** This reasoning might be especially problematic when it comes to demonstrating positive effects of mind wandering. In Manuscripts 3 and 4, we had expected beneficial mind wandering in the form of (associative and) continuous thought that produces (creative) solutions to previously presented problems during incubation-task intervals. One might argue that by continually inter- and disrupting the

---

[18] However, recent findings by Schubert et al. (2019) suggest that TUT reports' susceptibility to probe variations might not be a threat to thought-probes' validity. In their study, changes in probe characteristics were not reflected in changes in the relationship between mind wandering and SART performance. Thus, changes in TUT reports due to variations in thought probing methods might simply reflect varying absolute levels of reported mind wandering, possibly due to response biases. At least when held constant in an experimental setting (or when controlled for), such variations on the absolute level do not threaten the validity of probe-caught mind wandering.

ongoing incubation task, detrimental effects concerning the wandering mind's natural dynamics might occur. We therefore specified an *interruption hypothesis* in Manuscripts 3 and 4, which stated that thought probes' interrupting nature might cut off associative, continuous, and possibly beneficial trains of thought, influencing the mind-wandering experience. On the one hand, this interruption might not give the participants enough time to produce TUTs before being interrupted once again. On the other hand, probing on a consistent basis might be too disruptive, so that participants are prevented from focusing on the current task, producing even higher rates of TUTs (Seli, Carriere, et al., 2013).

We further specified an (alternative or) additional *awareness hypothesis*, which goes beyond purely interrupting effects. Since thought probes might increase meta-awareness and alert participants of their state of thought, participants might become more aware and cautious of their own mind-wandering behavior, possibly changing the experience itself, including potentially beneficial TUTs. On the one hand, participants could try to down-regulate their mind-wandering behavior because thought probes could make them believe that they should be thinking about the task. On the other hand, thought probes might remind participants of the opportunity to mind wander, leading them to engage in more TUTs. If both interruption and awareness play a role, there further might be a cumulative effect, causing even more changes in mind-wandering behavior. In Manuscripts 3 and 4, we expected such changes due to interruption and/or awareness to interfere with beneficial mind-wandering processes. We expected thought probes to disrupt creative and problem-solving TUTs.

To test and disentangle the proposed interruption and awareness hypotheses, we developed a new experimental paradigm, which we employed in Manuscripts 3 and 4. Three experimental conditions were crucial for this paradigm: In all conditions, participants were presented with a (creativity or apartment) problem and engaged in a problem-solution phase, followed by a filler interval (see Figure 3) and a subsequent second problem-engagement phase (see Figure 2). In the first (baseline) condition, participants worked on the filler task without any probes interrupting the task. In the second condition, participants worked on the same filler task and were interrupted by *trivia probes* from time to time, which consisted of general knowledge questions. Trivia probes should prompt an interruption, as proposed in the interruption hypothesis. If mere interruption were sufficient to evoke the proposed detrimental effects of thought probing, mind-wandering behavior (during the filler interval) for participants in this condition should deviate from the behavior observed in the baseline condition. If thought awareness were a necessary criterion for detrimental effects of thought probing (i.e., interruption by itself were not a sufficient disruption), trivia probes, in contrast to thought probes, should not have an effect on mind wandering. Consequently, in the third condition, including the to-be-tested awareness component, participants working on the filler task were interrupted by typical thought probes (same number and format as trivia probes). Finally, before working on the second problem-engagement phase, all participants filled in a retrospective questionnaire concerning on-task and off-task experiences during the incubation task. Retrospectively reported off-task experiences represented our central dependent variables. We further expected mind-wandering benefits, namely better performance in the second problem-engagement phase, to be greatest for participants in the uninterrupted baseline condition as their minds were allowed to actually run freely without any disturbance. Mind-wandering benefits were expected to depend on possible detrimental interruption and/or awareness effects, as trivia and/or thought probing may have disruptive effects on associative, continuous problem-solving processes during incubation.

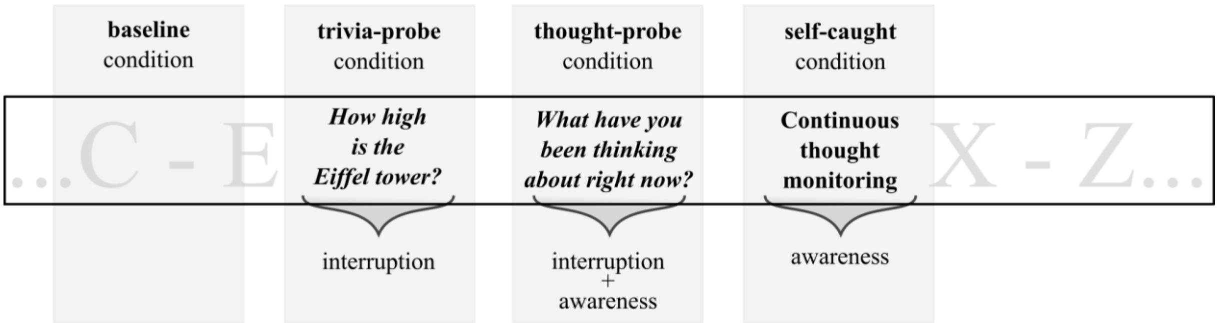


*Figure 3.* Schematic overview of the newly developed paradigm's filler interval. Participants in the baseline condition worked on the letter task without being interrupted and without instructions concerning the mind-wandering construct. Participants in the trivia- and thought-probe conditions were interrupted by the respective probes asking trivia questions or questions concerning current thoughts. Trivia probes were supposed to induce interruption effects, whereas thought probes were supposed to additionally induce awareness effects. Pure awareness effects were supposed to be induced within the self-caught condition (only employed in the unpublished study) by asking participants to continuously monitor their thoughts during the letter task.

In both manuscripts, we differentiated between general TUTs (unrelated to the current task and the to-be-solved active problem) and TUTs concerning the previously presented to-be-solved problem, as measured by the retrospective questionnaire. In Manuscript 3 (Experiment 3), we found that participants who had received thought probes reported significantly higher levels of general TUTs than participants who had received trivia probes or no thought probes. We did not find group differences for TUTs related to the apartment problem, which participants were supposed to solve. In Manuscript 4, we did not find group differences for general TUTs. However, there were significant group differences for TUTs related to the creativity problem. Participants in the thought-probe condition reported lower levels of creativity-problem-related thoughts than participants in both the baseline and the trivia-probe condition[19]. Figure 2 shows the result pattern for problem-related TUTs for Manuscripts 3 and 4 as well as an additional yet unpublished study, replicating Manuscript 4.

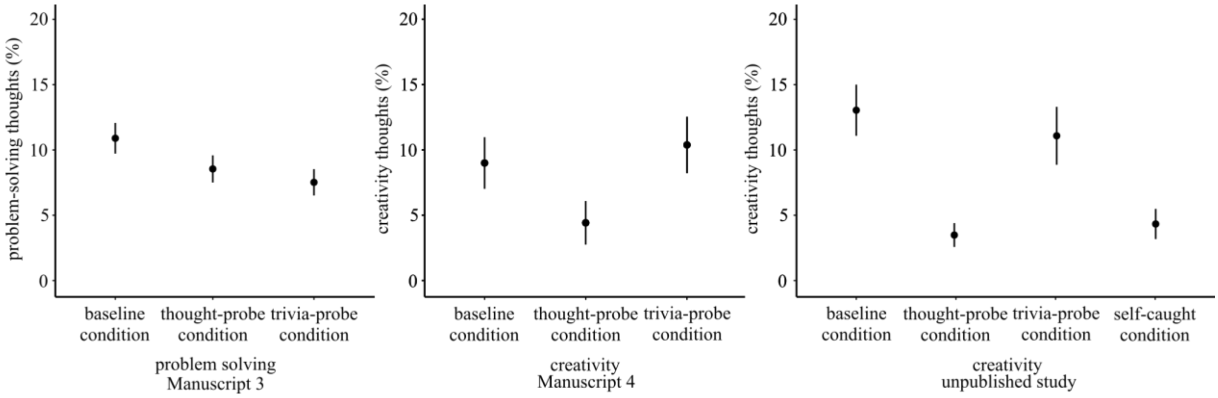


*Figure 4.* Depiction of rates of problem-related mind wandering collected via the new paradigm as reported in Manuscripts 3 and 4 as well as in an unpublished study. In Manuscript 3, apartment-problem-related thoughts did not vary between conditions. In Manuscript 4 and in the unpublished study, awareness-inducing manipulations (see thought probe conditions and self-caught condition) yielded lower reports of creativity-problem-related thoughts as the baseline and only-interruption inducing conditions.

In the unpublished study (a preregistration can be found at https://doi.org/10.17605/OSF.IO/X74F2), I added a fourth condition to the new paradigm, in which participants were not interrupted by any kinds of probes during the incubation phase between two creativity problem-solving phases[20]. However, before incubation they were asked to monitor their own on- and off-task thought behavior and to press a button whenever they caught themselves mind wandering (self-caught mind wandering assessment, e.g., Weinstein, 2018). After incubation, they answered to the same retrospective thought questionnaire (see above). This fourth condition can be interpreted as a *pure awareness* condition. Such a condition usefully complemented the paradigm employed in Manuscripts 3 and 4, as there were *pure interruption* and *combined interruption and awareness*, but no pure

---

[19] For a more detailed description of the experiments' materials and results, please see Appendices A3 and A4.
[20] Participants worked on two Unusual Consequences Tasks (Fulgosi & Guilford, 1968; Hass & Beaty, 2018).

awareness conditions. As apparent from Figure 2, results from this unpublished study replicated results from Manuscript 4 concerning creativity TUTs and further make a point in favor of the awareness hypothesis: Participants in the thought-probe and the self-caught conditions (both including an awareness component) showed lower levels of thoughts regarding the creativity task compared to participants in the baseline and the trivia-probe conditions (both excluding an awareness component). In the unpublished study, as in Manuscript 4, there were no group differences concerning general TUTs. In all studies, however, group differences in mind-wandering behavior were unrelated to apartment-task or creativity performance. As discussed in sections 2.3.2 and 2.3.3, we did not find beneficial mind-wandering effects within the problem-solving and creativity domains. We could thus not relate disruptive effects of thought probes to the non-appearance of mind-wandering benefits.

**3.1.2 Discussion of the New Paradigm.** In Manuscripts 3 and 4 as well as in a yet unpublished study, we developed and tested a new paradigm to disentangle possible interruption and awareness effects of thought probes. In all studies, thought probes had an effect on retrospectively reported mind-wandering behavior, although in Manuscript 3 this pertained to general TUTs whereas in Manuscript 4 as well as in the unpublished study this pertained to problem-related (creativity) TUTs. It has, however, yet to be determined why different to-be-solved problems produce different effects thought-probe existence.

Participants who had been interrupted by trivia probes generally reported similar levels of TUTs (general as well as problem-related) as participants in the baseline conditions. It seems that the interrupting nature of thought probes is not disruptive to thought experiences during incubation. However, thought probes as well as self-caught instructions led to differences in incubation mind-wandering, indicating that thought awareness might influence actually experienced or reported thought processes. Especially the combination of results from Manuscript 4 and the unpublished study speaks for the interfering nature of thought awareness when it comes to problem-related (creative) thought processes. Levels of creativity TUTs were roughly halved in awareness-inducing experimental conditions[21]. It remains an open question, which mechanisms underlie this reduction in creative thoughts. Thought awareness might change actual mind wandering by altering the participants of their thought-processes, leading to a suppression of TUTs. Thought awareness might, however, also change

---

[21] It should, however, be mentioned that this lack of creative thought processes was not reflected in worse post-incubation creativity performance.

later reports of mind wandering due to demand characteristics without actually changing mind-wandering behavior: Participants might have felt that it was inappropriate to mind wander creatively because thought awareness might have made them speculate about the background of the experiment. Independent of the underlying mechanisms, thought probes seem to make it difficult to detect *true* mind-wandering rates.

Thought probes have been found to not influence current task performance (Wiemers & Redick, 2019). However, to our knowledge, we were the first to show that such probes influence current mind-wandering behavior or reports. We also took a closer look at the processes which might underlie this TUT alteration, namely interruption as well as awareness effects. As thought probes are used highly frequently in mind-wandering research (section 1.2), it is of great importance to the field that we as researchers know how this tool works and what effects it might evoke. In Manuscript 4 as well as in the unpublished study, thought probes decreased creative TUTs (or reports of the latter), in other words, thoughts directed at a pending problem. Representing a limitation of our results, we did not find a negative relation between such a TUT decrease and beneficial mind-wandering effects. We had expected the latter to occur as long as minds were allowed to wander freely without any interruption. In the domains of creativity and problem solving, however, we did not find any mind-wandering benefits independent of probing conditions. This does not rule out the possibility that thought probes might impede the search for positive effects of mind wandering in other domains in which possibly productive TUTs might be affected by thought probing. It might, for example, be difficult to demonstrate positive effects of mind-wandering episodes as opportunities for mental breaks or as having a dishabituating function (see section 2.1 and Schooler et al., 2011) when people are constantly made aware of their own thought processes. Further, in Manuscript 2, mind-wandering benefits might have been more substantial if we had refrained from making participants aware of their memory-maintenance behavior due to thought probing. Further research should focus on these assumptions and identify further domains as well as types of TUTs which are influenced by the common method of thought probing.

## 4 Eye-Movement Measures as Mind-Wandering Indicators

### *4.1 Objective Mind-Wandering Assessment*

The apparent subjectivity of thought probes as well as the previously described problems (section 3), which come with thought-probe employment, have led to a search for objective biological or behavioral indicators of TUTs. The central aims behind this search are the validation of subjective thought reports, but further also the direct measurement of mind wandering without a subjective component. The basic idea behind such direct and objective measurement methods is, that when thoughts trail off, the external perception is affected. Put more precisely, during mind wandering, attention is decoupled from external, perceptual input towards internal processing (Schooler et al., 2011; Stawarczyk et al., 2011). Objective mind-wandering markers are supposed to make such an *attentional decoupling* process visible and several different markers have been proposed.

*Reaction times* were found to be sensitive to attentional lapses with people showing an increased variability on this measure when mind wandering (McVay & Kane, 2012a; Seli, Cheyne, et al., 2013). In addition, *neurological markers* are being explored. The *Default Mode Network* was identified as a set of brain regions becoming active when the brain is at rest. However, it is not a passive network, as it was found to play an active role in internally directed cognition. Its activation was connected to mind wandering in several studies and seems to be a stable mind-wandering correlate (Andrews-Hanna et al., 2014; Buckner et al., 2008; Christoff et al., 2009; Mason et al., 2007; Mittner et al., 2016). As a further neurological marker for drifting thoughts, the *P300 event-related potential* (an electroencephalography (EEG) measure) was proposed, as it was found to be reduced prior to TUT episodes (Smallwood, Beach, et al., 2008). Not only the brain, but also the eyes seem to reveal the wandering mind. *Oculomotor behaviors* such as *pupil dilation* and *eye movements* can be temporally associated with TUTs. Pupillary responses have not only been found to distinguish between on-task and off-task states (however with disagreement on the direction of effects, see Franklin, Broadway, et al., 2013; Grandchamp et al., 2014; Smallwood et al., 2011), but Unsworth and Robison (2016) even suggested that they are sensitive to different types of inattention triggered by internal (mind wandering) versus external distraction.

### *4.2 Eye Movements During Mindless Reading (Manuscript 1[22])*

Eye movements, especially those during reading[23], have also been examined as potential biological markers of mind-wandering processes. The objectivity of such measures

[22] Steindorf, L., & Rummel, J. (2020). Do your eyes give you away? A validation study of eye-movement measures used as indicators for mindless reading. *Behavior Research Methods, 52*(1), 162-176. https://doi.org/10.3758/s13428-019-01214-4

[23] The question of how mind wandering affects eye movements during reading has already been addressed in several studies. However, there are yet other domains in which eye movements have been found to be sensitive towards a drifting mind, such as driving (He et al., 2011), video lectures (Zhang et al., 2018), or movies (Mills et al., 2016).

is given because people are not able to deliberately control fixations and saccades[24] (Rayner, 1998). Also, from a practical perspective, the method of eye tracking has become more and more feasible in the last decades with easily usable devices and software available on the market (Holmqvist et al., 2011). Thus, would changes in eye movements robustly correlate with the occurrence of mind-wandering episodes, these measures could represent a suitable candidate for non-subjective mind-wandering assessment. The reason why eye movements might be associated with attentional decoupling (see above) and therefore the motivation to identify them as biological markers for mind-wandering episodes lies in the general idea behind eye-tracking research: There is a link between what people are looking at and what they are currently mentally processing (Uzzaman & Joordens, 2011). Given an attentional-decoupling process during mind wandering, this link should be interrupted. To transfer this logic to the domain of reading, the duration of a fixation on a given word partially reflects the duration of its processing. The word-frequency effect (Inhoff & Rayner, 1986) nicely illustrates this idea: People take longer to identify and process an uncommon than a common word. This need for more intense processing is reflected in longer fixations on uncommon words. Thus, one could argue that cognitive forces drive eye-movement behavior (*cognitive-control theories*, see Just & Carpenter, 1980; but for o*culomotor-control theories* see, Yang, 2006). However, during mind wandering, these cognitive forces are engaged in internally generated thoughts, which might limit their power to control eye movements. Consequently, when thoughts (and therefore attention) drift away from lines of text, eye movements should become less sensitive towards lexical features, such as word frequency.

Which specific eye-movement measures are affected by a wandering mind was the central research question of several previous studies, all of which comparing eye movements during episodes of normal in comparison to *mindless* reading (Foulsham et al., 2013; Frank et al., 2015; Reichle et al., 2010; Smilek et al., 2010; Uzzaman & Joordens, 2011). Reading modes were self-classified by the participants using thought probes. This way, objective eye-movements were supposed to validate subjective mind-wandering self-reports (and vice versa). In Manuscript 1, we reviewed these studies and found little consensus regarding the specific eye-movement measures that were found to be sensitive to attentional decoupling. As the search for eye-movement indicators of a wandering mind still seems to be a work in

---

[24] Eyes do not move continuously and “smoothly”. Simply put, almost all eye movements can be considered a sequence of fixations (short stops to “gather” information) and saccades (short, rapid movements to the next stop).

progress, we conducted a high-powered eye-tracking experiment with the aim to validate all previously identified indicators.

We asked participants to read a text for up to 30 minutes with their eye movements being tracked. Reading was self-paced, and participants were asked about their current thoughts after having read each of ten target sentences. These target sentences had been determined before data collection and five of them contained a high-frequency and the other five a low-frequency target word. After reading, a retrospective mind-wandering questionnaire as well as a reading-comprehension test followed. We were interested in differences in eye-movement measures between self-classified normal- and mindless-reading episodes. These comparisons regarded target sentences as well as target words. Concerning target sentences, the previous studies we had reviewed had found mixed evidence regarding six different possible mind-wandering indicators. For only three of them (reading time, fixation count, first-fixation duration), we found significant differences between on-task and off-task episodes. All of these three measures can be described as reading-speed related. When reading mindlessly, our participants slowed down, exhibiting longer and more fixations on the target sentences. Consequently, it took them longer to mindlessly “read” a sentence. Concerning target words, we compared the magnitudes of the word-frequency effect for normal- and mindless-reading episodes. For the fixation-count measure (number of fixations on a given word), we found smaller differences between high- and low-frequency target words during off-task episodes, i.e. a smaller word-frequency effect. Word features like word frequency seem to have a smaller impact on eye movements when people’s minds start to wander than they usually have during reading, which is well in line with the idea of attentional decoupling[25].

Identifying objective measures of mind wandering is an important endeavor. However, in Manuscript 1, we found that it is still a work in progress. Our study constitutes one among several others assessing eye-movements to detect mind-wandering indicators, and it is noticeable that the studies’ results differ. Thus, it was all the more important for our study to review and combine the preexistent work, aiming at further validating previously reported effects. Although not all effects could be replicated, reading-speed-related measures turned out to be promising candidates for mind-wandering detection via eye-tracking methods. Identifying online-markers of mind wandering is not only important because it contributes to

[25] For a more detailed description of the experiment’s materials and results, please see Appendix A1.

theoretical assumptions of mind wandering and attentional decoupling. It could also eventually contribute to a non-intrusive, objective mind-wandering measurement, circumventing the intrusive effects of thought probes and the bias-prone nature of self-report measures. Further, from an applied perspective, objective indicators might foster the development of intervention systems for situations in which mind wandering is disruptive (e.g., educational context) and even dangerous (e.g., while driving). Still, more work is necessary, and we hope for our work to drive the validation process forward. The ultimate goal should be to develop algorithms that allow for real-time detection of mind wandering during reading, based on previously validated eye-movement measures.

## 5 General Discussion and Conclusion

### *5.1 Good Thoughts, Bad Thoughts?*

This dissertation investigated the wandering mind's nature. I contrasted previously reported negative consequences of mind wandering ("bad thoughts") with positive ones ("good thoughts"), before reporting new contributions to this line of research. The functionality and valence of mind-wandering processes appeared to be rather complex, depending, in the first place, on context and content. Section 2.2 focused on context-related issues, stating that a mind-wandering episode's outcome-valence is situation- and trait-specific. There are situations in which mind wandering is appropriate and even desirable, such as low-difficulty tasks. In other situations (e.g., high-difficulty tasks), mind wandering should be inhibited. The same off-task thoughts might thus be *good* in one situation, but *bad* in the other. There are also person-related factors which control or benefit adjustment processes in accordance with current situational demands (e.g., working memory), suggesting individual differences when it comes to mind-wandering outcome-valence.

The discussion of section 2.3 covered content-related issues on which TUTs' outcome valence might depend. Some mind-wandering content might be *bad* (or at least *not* good or useful), whereas other content might be *good*. I argued that personally relevant and goal-oriented prospective TUTs (as in Manuscript 2) rank among the *good*, and that a high concreteness and simplicity of task formulation and task goals also facilitate the usefulness of TUT episodes, at least within laboratory experiments. Such content-related factors could explain why we found mind-wandering benefits within the domain of memory, but not within the domains of creativity or problem solving.

I further argued (section 2.3.4) that benefits of mind wandering might reveal themselves under different circumstances than costs might do and that different methodological approaches could be necessary when it comes to detecting *good* versus *bad* thoughts. When examining mind-wandering benefits, thought content becomes crucial and the researcher should not only refer to mind-wandering frequencies, as is often the case when examining costs. A further difference between costs and benefits might be that costs often reveal themselves right away whereas benefits might occur at a later time, also speaking for the use of different methodological approaches. I further considered the question of whether *good* and *bad* is equal for everyone by suggesting an individual-differences view on mind-wandering benefits.

To summarize, the question of *good* or *bad* is not an easy one to answer and demands a close investigation of situation- and person-related factors. Mind wandering should not be viewed one-sidedly, because, besides apparent and exhaustively studied costs, benefits exist. I added memory to the list of domains which have been found to benefit from TUT episodes and discussed moderating factors which should be further examined in future research, as described in more detail in the respective sections. Moreover, I would encourage the search for additional positive functions of mind-wandering processes (section 2.1, e.g., *relief from boredom*, *mental rest*, *emotion regulation*) to eventually complete a balanced picture of the wandering mind's nature.

***5.2 Can We Measure Mind Wandering Online?***

This dissertation further focused on how to capture wandering thoughts. In section 1.2, subjective in-the-moment mind-wandering assessment methods were introduced, and I stated reasons for thought probes' frequent usage within mind-wandering research. These reasons were contrasted with certain deleterious effects they evoke (section 3). As thought probes might influence thought processes by interrupting participants and making them aware of their current thoughts, we developed a new paradigm to investigate such effects (section 3.1.1). We found that interruption does not affect participant's retrospectively reported thoughts. Awareness, however, reduced possibly creative thoughts in two experiments. Thought probes had already been found to not influence current task performance (Wiemers & Redick, 2019). We, however, now showed that they do influence thought processes or at least retrospective thought reports. Such intruding effects as well as the apparent subjectivity of self-report measures have led to a search for objective markers of TUTs (section 4.1), and we examined eye-movement measures as possible mind-wandering indicators, finding speed-related variables to be associated with mind-wandering episodes.

As of now, no *gold-standard* for measuring in-the-moment mind wandering exists. Self-reports are subjective and some of them are intruding; identification of objective markers is yet a work in progress. For relative group comparisons such as those between experimental conditions, subjective measures should suffice, as long as their distorting effects are stable (Schubert et al, 2019). For absolute measures of mind-wandering levels or for the investigation of relations between mind wandering and other factors (such as creativity performance, for example), they might be inadequate. Future research should thus focus on developing reliable objective, biological markers of TUTs. EEG as well as eye-movement and pupillometry markers are promising candidates because of their close connection to attentional processes and their high temporal resolutions (section 4). However, software, hardware, as well as corresponding algorithms linking such markers to mind-wandering behavior (e.g., Franklin et all., 2011) are still in need of improvement.

Further, it may be problematic that objective measures are most often validated using only (arguably flawed and biased) subjective measures. This limitation also concerns Manuscript 1. As of now, lacking a *gold-standard* measure, this seems to be the common practice when it comes to the development of objective assessment methods. In the future, however, validation studies should at best include more than one objective marker, as well as external variables from various domains that have previously been connected to mind-wandering processes (e.g., personality, working memory, etc.). Including newly developed objective markers within such a wide framework would allow for stronger validity testing.

### *5.3 Practical Implications*[26]

Mind Wandering often represents a source of disturbance and error (section 1.3). One might thus easily get the impression that it should be inhibited at all times as it hampers focused work. However, a balanced view on the mind-wandering phenomenon and especially the growing list of mind-wandering benefits (section 2) implicate that besides intervention

---

[26] Besides practical implications, this dissertation offers methodological ones. However, as they have already been covered, they will only be shortly listed here: Considering adaptive mind wandering from a methodological perspective (section 2.3.4), I suggested that (1) focusing on thought content (e.g., via assessing thought descriptions) rather than thought frequency should be crucial when it comes to detecting beneficial effects. I further (2) argued that methods aiming at revealing positive mind-wandering outcomes should consider time-delayed effects (e.g., via long-term assessments using mobile-phone applications). Moreover, I (3) discussed adaptive mind wandering within an individual-differences approach and recommended including trait-based moderators when examining mind-wandering benefits. Considering methodological issues, I highlighted the importance of knowledge concerning employed research tools (section 3.1.2) and (4) suggested that thought probes can be a suitable choice of method (section 5.2) but should be applied with caution. Finally (5), eye-movements might represent a promising measure to detect TUT episodes, but we showed that more validation work is necessary (section 4).

strategies that focus solely on decreasing mind-wandering frequencies in daily life (e.g., Bennike et al., 2017; Mrazek et al., 2012; Rahl et al., 2017), strategies focusing on regulation processes could be of use. For example, within work and educational contexts, it could be helpful for people to know when to wander and when not to wander (section 2.2) and to regulate thoughts accordingly to minimize costs and maximize benefits. Some people seem to be able to detect situations in which mind wandering is inappropriate and to downregulate TUTs accordingly (Rummel & Boywitt, 2014). On the other hand, in situations in which mind wandering is appropriate or even desirable, they show higher TUT levels. People who are not intuitively able to do so could be supported in their use of such regulation strategies. First steps might be creating awareness for wandering thoughts as well as for current task requirements and explicitly categorizing self-caught mind wandering as good or bad. For good mind wandering, further strategies or interventions should focus on shifting abilities, for bad mind wandering on inhibition abilities (Miyake & Friedman, 2012).

In everyday life, some people might suffer more from wandering thoughts than others. For example, current research suggests that TUTs, especially those arising spontaneously, represent a central feature of ADHD symptomatology (Bozhilova et al., 2018; Seli, Smallwood, et al., 2015b). ADHD patients might benefit particularly well from intervention methods that focus on the detection of wandering thoughts in order to further inhibit them when they are inappropriate. However, as research employing self-caught methods (section 1.2) unfolds, sometimes wandering thoughts are hard to detect (Schooler, 2002; Schooler et al., 2004), maybe even more so for people experiencing them highly frequently on a daily basis. Objective mind-wandering markers (when improved and practicable, section 5.2), especially ones that are easy to assess such as eye-movements (section 4) or pupil sizes, could facilitate the detection of wandering thoughts, for example while reading or studying. Within training sessions, (not only) ADHD patients could be informed as soon as their thoughts start drifting, giving them a feeling for their own thought processes and enabling them to explicitly work on inhibition strategies.

Leaving the educational context, detection and inhibition strategies are further especially desirable in situations where mind wandering is dangerous. As objective, biological markers (section 4) could bring us one step closer to real-time detection, they might have the potential to increase road safety. Driving simulation studies have shown that wandering thoughts can quickly become dangerous (Albert et al., 2018; Lin et al., 2016; Yanko & Spalek, 2014) due to, for example, increasing braking reaction times. Eye-state bio-feedback methods to detect *fatigue*, a further cause of traffic accidents, have already been tested (e.g., Devi & Bajaj, 2008). Further eye-focused methods detecting attentional lapses could warn drivers when their thoughts are drifting off and could thus be in the interest of road safety.

### *5.5 Conclusion*

This dissertation aimed at painting a balanced picture of the wandering mind's nature and at investigating subjective as well as objective measurement methods. Costs and benefits of mind wandering were reported and brought in line. Memory was introduced as a newly considered domain found to benefit from mind-wandering processes, before possible moderators of *good versus bad* were discussed. As mind-wandering research is dependent on reliable thought-measures, first subjective and then objective assessment methods were investigated and discussed.

As a relatively new research field, which however "exploded" starting 2006 (Callard et al., 2013; Weinstein, 2018), many aspects of mind wandering have been extensively studied whereas there is still much to be discovered. Furthermore, there is already some "tidying up" to do (see, for example, Manuscript 1), as similar strings of research are developing in parallel and researchers are generating new methods on an almost daily basis (Weinstein, 2018). Besides mind wandering's ubiquity in everyday life and one's own personal experiences with the topic, this only makes the research field all the more interesting and worthwhile to be studied further.

**References**

Acker, F. (2008). New findings on unconscious versus conscious thought in decision making: Additional empirical data and meta-analysis. *Judgment and Decision Making*, *3*(4), 292-303. http://journal.sjdm.org/jdm71128.pdf

Albert, D. A., Ouimet, M. C., Jarret, J., Cloutier, M.-S., Paquette, M., Badeau, N., & Brown, T. G. (2018). Linking mind wandering tendency to risky driving in young male drivers. *Accident Analysis & Prevention*, *111*, 125-132. https://doi.org/10.1016/j.aap.2017.11.019

Andrews-Hanna, J. R., Smallwood, J., & Spreng, R. N. (2014). The default network and self-generated thought: Component processes, dynamic control, and clinical relevance. *Annals of the New York Academy of Sciences*, *1316*, 29-52. https://doi.org/10.1111/nyas.12360

Baars, B. J. (2010). Spontaneous repetitive thoughts can be adaptive: Postscript on "mind wandering". *Psychological Bulletin*, *136*(2), 208-210. https://doi.org/10.1037/a0018726

Baird, B., Smallwood, J., Mrazek, M. D., Kam, J. W. Y., Franklin, M. S., & Schooler, J. W. (2012). Inspired by distraction: Mind wandering facilitates creative incubation. *Psychological Science*, *23*(10), 1117-1122. 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. *Conscious Cognition*, *20*(4), 1604-1611. https://doi.org/10.1016/j.concog.2011.08.007

Bennike, I. H., Wieghorst, A., & Kirk, U. (2017). Online-based mindfulness training reduces behavioral markers of mind wandering. *Journal of Cognitive Enhancement*, *1*(2), 172-181. https://doi.org/10.1007/s41465-017-0020-9

Bozhilova, N. S., Michelini, G., Kuntsi, J., & Asherson, P. (2018). Mind wandering perspective on attention-deficit/hyperactivity disorder. *Neuroscience & Biobehavioral Reviews, 92*, 464-476. https://doi.org/10.1016/j.neubiorev.2018.07.010

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

Callard, F., Smallwood, J., Golchert, J., & Margulies, D. (2013). The era of the wandering mind? Twenty-first century research on self-generated mental activity. *Frontiers in Psychology*, *4*, Article 891. https://doi.org/10.3389/fpsyg.2013.00891

Calvillo, D. P., & Penaloza, A. (2009). Are complex decisions better left to the unconscious? Further failed replications of the deliberation-without-attention effect. *Judgment and Decision Making*, *4*(6), 509-517. http://journal.sjdm.org/9711/jdm9711.pdf

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, *132*(3), 354-380. https://doi.org/10.1037/0033-2909.132.3.354

Cheyne, A. J., Solman, G. J. F., Carriere, J. S. A., & Smilek, D. (2009). Anatomy of an error: A bidirectional state model of task engagement/disengagement and attention-related errors. *Cognition*, *111*(1), 98-113. https://doi.org/10.1016/j.cognition.2008.12.009

Christoff, K., Gordon, A. M., Smallwood, J., Smith, R., & Schooler, J. W. (2009). Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proceedings of the National Academy of Sciences*, *106*(21), 8719-8724. https://doi.org/10.1073/pnas.0900234106

Christoff, K., Irving, Z. C., Fox, K. C. R., Spreng, R. N., & Andrews-Hanna, J. R. (2016). Mind-wandering as spontaneous thought: A dynamic framework. *Nature Reviews Neuroscience*, *17*(11), 718-731. https://doi.org/10.1038/nrn.2016.113

Csifcsák, G., & Mittner, M. (2017). Linking brain networks and behavioral variability to different types of mind-wandering. *Proceedings of the National Academy of Sciences*, *114*(30), E6031-E6032. https://doi.org/10.1073/pnas.1705108114

D'Argembeau, A., Renaud, O., & Van Der Linden, M. (2011). Frequency, characteristics and functions of future-oriented thoughts in daily life. *Applied Cognitive Psychology*, *25*(1), 96-103. https://doi.org/10.1002/acp.1647

Devi, M. S., & Bajaj, P. R. (2008). *Driver fatigue detection based on eye tracking*. Paper presented at the 2008 First International Conference on Emerging Trends in Engineering and Technology, Nagpur, India.

Dijksterhuis, A., Bos, M. W., Nordgren, L. F., & van Baaren, R. B. (2006). On making the right choice: The deliberation-without-attention effect. *Science*, *311*(5763), 1005-1007. https://doi.org/10.1126/science.1121629

Dijksterhuis, A., & Nordgren, L. F. (2006). A theory of unconscious thought. *Perspectives on Psychological Science*, *1*(2), 95-109. https://doi.org/10.1111/j.1745-6916.2006.00007.x

Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology*, *84*(5), 795-805. https://doi.org/10.1037/0021-9010.84.5.795

Dweck, C. S., Chiu, C.-y., & Hong, Y.-y. (1995). Implicit theories and their role in judgments and reactions: A word from two perspectives. *Psychological Inquiry*, *6*(4), 267-285. https://doi.org/10.1207/s15327965pli0604_1

Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology* (H. A. Ruger & C. E. Bussenius, Trans.). Dover Publications. (Original work published 1885)

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*(3), 215-251. https://doi.org/10.1037/0033-295X.87.3.215

Farley, J., Risko, E., & Kingstone, A. (2013). Everyday attention and lecture retention: The effects of time, fidgeting, and mind wandering. *Frontiers in Psychology*, *4*, Article 619. https://doi.org/10.3389/fpsyg.2013.00619

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

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). McGraw-Hill Book Company.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, *34*(10), 906-911. https://doi.org/10.1037/0003-066X.34.10.906

Foulsham, T., Farley, J., & Kingstone, A. (2013). Mind wandering in sentence reading: Decoupling the link between mind and eye. *Canadian Journal of Experimental Psychology*, *67*(1), 51-59. https://doi.org/10.1037/a0030217

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

Franklin, M. S., Broadway, J. M., Mrazek, M. D., Smallwood, J., & Schooler, J. W. (2013). Window to the wandering mind: Pupillometry of spontaneous thought while reading. *The Quarterly Journal of Experimental Psychology*, *66*(12), 2289-2294. https://doi.org/10.1080/17470218.2013.858170

Franklin, M. S., Mrazek, M., Anderson, C., Smallwood, J., Kingstone, A., & Schooler, J. (2013). The silver lining of a mind in the clouds: Interesting musings are associated with positive mood while mind-wandering. *Frontiers in Psychology*, *4*, Article 583. https://doi.org/10.3389/fpsyg.2013.00583

Franklin, M. S., Smallwood, J., & Schooler, J. W. (2011). Catching the mind in flight: Using behavioral indices to detect mindless reading in real time. *Psychonomic Bulletin & Review*, *18*(5), 992-997. https://doi.org/:10.3758/s13423-011-0109-6

Fulgosi, A., & Guilford, J. P. (1968). Short-term incubation in divergent production. *The American Journal of Psychology*, *81*(2), 241-246. https://doi.org/10.2307/1421269

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

Godbole, N. R., Delaney, P. F., & Verkoeijen, P. P. J. L. (2014). The spacing effect in immediate and delayed free recall. *Memory*, *22*(5), 462-469. https://doi.org/10.1080/09658211.2013.798416

Grandchamp, R., Braboszcz, C., & Delorme, A. (2014). Oculometric variations during mind wandering. *Frontiers in Psychology*, *5*, Article 31. https://doi.org/10.3389/fpsyg.2014.00031

Hao, N., Ku, Y., Liu, M., Hu, Y., Grabner, R. H., & Fink, A. (2014). Enhancing verbal creativity via brief interventions during an incubation interval. *Creativity Research Journal*, *26*(1), 30-38. https://doi.org/10.1080/10400419.2014.873658

Hass, R. W., & Beaty, R. E. (2018). Use or consequences: Probing the cognitive difference between two measures of divergent thinking. *Frontiers in Psychology*, *9*, Article 2327. https://doi.org/10.3389/fpsyg.2018.02327

He, J., Becic, E., Lee, Y.-C., & McCarley, J. S. (2011). Mind wandering behind the wheel: Performance and oculomotor correlates. *Human Factors*, *53*(1), 13-21. https://doi.org/10.1177/0018720810391530

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures* (1st ed.). Oxford University Press.

Huba, G. J., Aneshensel, C. S., & Singer, J. L. (1981). Development of scales for three second-order factors of inner experience. *Multivariate Behavioral Research*, *16*(2), 181-206. https://doi.org/10.1207/s15327906mbr1602_4

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics*, *40*(6), 431-439. https://doi.org/10.3758/bf03208203

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*(4), 329-354. https://doi.org/10.1037/0033-295X.87.4.329

Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science*, *330*(6006), 932-932. https://doi.org/10.1126/science.1192439

Klinger, E. (1967). Modeling effects on achievement imagery. *Journal of Personality and Social Psychology*, *7*(1), 49-62. https://doi.org/10.1037/h0024936

Klinger, E. (1999). Thought flow: Properties and mechanisms underlying shifts in content. In J. A. Singer & P. Salovey (Eds.), *At play in the fields of consciousness: Essays in honor of Jerome L. Singer* (pp. 29-50). Lawrence Erlbaum Associates Publishers.

Klinger, E. (2013). Goal commitments and the content of thoughts and dreams: Basic principles. *Frontiers in Psychology*, *4*, Aricle 415. https://doi.org/10.3389/fpsyg.2013.00415

Kong, D. T., & Jolly, P. M. (2019). Lay beliefs about attention to and awareness of the present: Implicit mindfulness theory (IMT) and its workplace implications. *Journal of Business and Psychology*, *34*(5), 685-707. https://doi.org/10.1007/s10869-018-9606-1

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, *103*(3), 490-517. https://doi.org/10.1037/0033-295X.103.3.490

Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences*, *36*(6), 661-679. https://doi.org/10.1017/S0140525X12003196

Leszczynski, M., Chaieb, L., Reber, T. P., Derner, M., Axmacher, N., & Fell, J. (2017). Mind wandering simultaneously prolongs reactions and promotes creative incubation. *Scientific Reports*, *7*, Article 10197. https://doi.org/10.1038/s41598-017-10616-3

Lin, C.-T., Chuang, C.-H., Kerick, S., Mullen, T., Jung, T.-P., Ko, L.-W., Chen, S.-A., King, J.-T., & McDowell, K. (2016). Mind-wandering tends to occur under low perceptual demands during driving. *Scientific Reports*, *6*, Article 21353. https://doi.org/10.1038/srep21353

Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., & Macrae, C. N. (2007). Wandering minds: The default network and stimulus-independent thought. *Science*, *315*(5810), 393-395. https://doi.org/10.1126/science.1131295

McMillan, R. L., Kaufman, S. B., & Singer, J. L. (2013). Ode to positive constructive daydreaming. *Frontiers in Psychology*, *4*, Article 626. https://doi.org/10.3389/fpsyg.2013.00626

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

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

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

McVay, J. C., & Kane, M. J. (2012b). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of Experimental Psychology: General*, *141*(2), 302-320. https://doi.org/10.1037/a0025250

McVay, J. C., Kane, M. J., & Kwapil, T. R. (2009). Tracking the train of thought from the laboratory into everyday life: An experience-sampling study of mind wandering across controlled and ecological contexts. *Psychonomic Bulletin & Review*, *16*(15), 857-863. https://doi.org/10.3758/pbr.16.5.857

Mills, C., Bixler, R., Wang, X., & D’Mello, S. (2016). Automatic gaze-based detection of mind wandering during film viewing. In T. Barnes, M. Chi, & M. Feng (Eds.), *Proceedings of the 9th International Conference on Educational Data Mining 2016* (pp. 30–37). International Educational Data Mining Society.

Mittner, M., Hawkins, G. E., Boekel, W., & Forstmann, B. U. (2016). A neural model of mind wandering. *Trends in Cognitive Sciences*, *20*(8), 570-578. https://doi.org/10.1016/j.tics.2016.06.004

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*(1), 8-14. https://doi.org/10.1177/0963721411429458

Mooneyham, B. W., & Schooler, J. W. (2013). The costs and benefits of mind-wandering: A review. *Canadian Journal of Experimental Psychology*, *67*(1), 11-18. https://doi.org/10.1037/a0031569

Mrazek, M. D., Smallwood, J., Franklin, M. S., Chin, J. M., Baird, B., & Schooler, J. W. (2012). The role of mind-wandering in measurements of general aptitude. *Journal of Experimental Psychology: General*, *141*(4), 788-798. https://doi.org/10.1037/a0027968

Mrazek, M. D., Smallwood, J., & Schooler, J. W. (2012). Mindfulness and mind-wandering: Finding convergence through opposing constructs. *Emotion*, *12*(3), 442-448. https://doi.org/10.1037/a0026678

Newell, B. R., & Rakow, T. (2011). Revising beliefs about the merit of unconscious thought: Evidence in favor of the null hypothesis. *Social Cognition*, *29*(6), 711-726. https://doi.org/10.1521/soco.2011.29.6.711

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*(3), 231-259. https://doi.org/10.1037/0033-295X.84.3.231

Poerio, G. L., Totterdell, P., Emerson, L.-M., & Miles, E. (2016). Social daydreaming and adjustment: An experience-sampling study of socio-emotional adaptation during a life transition. *Frontiers in Psychology*, *7*, Article 13. https://doi.org/10.3389/fpsyg.2016.00013

Poh, J.-H., Chong, P. L. H., & Chee, M. W. L. (2016). Sleepless night, restless mind: Effects of sleep deprivation on mind wandering. *Journal of Experimental Psychology: General*, *145*(10), 1312-1318. https://doi.org/10.1037/xge0000207

Prebble, S. C., Addis, D. R., & Tippett, L. J. (2013). Autobiographical memory and sense of self. *Psychological Bulletin*, *139*(4), 815-840. https://doi.org/10.1037/a0030146

Preiss, D. D., Cosmelli, D., Grau, V., & Ortiz, D. (2016). Examining the influence of mind wandering and metacognition on creativity in university and vocational students. *Learning and Individual Differences*, *51*, 417-426. https://doi.org/10.1016/j.lindif.2016.07.010

Rahl, H. A., Lindsay, E. K., Pacilio, L. E., Brown, K. W., & Creswell, J. D. (2017). Brief mindfulness meditation training reduces mind wandering: The critical role of acceptance. *Emotion*, *17*(2), 224-230. https://doi.org/10.1037/emo0000250

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372-422. https://doi.org/10.1037/0033-2909.124.3.372

Reichle, E. D., Reineberg, A. E., & Schooler, J. W. (2010). Eye movements during mindless reading. *Psychological Science*, *21*(9), 1300-1310. https://doi.org/10.1177/0956797610378686

Rey, A., Goldstein, R. M., & Perruchet, P. (2009). Does unconscious thought improve complex decision making? *Psychological Research*, *73*(3), 372-379. https://doi.org/10.1007/s00426-008-0156-4

Robison, M. K., Miller, A. L., & Unsworth, N. (2019). Examining the effects of probe frequency, response options, and framing within the thought-probe method. *Behavior Research Methods*, *51*(1), 398-408. https://doi.org/10.3758/s13428-019-01212-6

Ruby, F. J. M., Smallwood, J., Sackur, J., & Singer, T. (2013). Is self-generated thought a means of social problem solving? *Frontiers in Psychology*, *4*, Article 962. https://doi.org/10.3389/fpsyg.2013.00962

Rummel, J., & Boywitt, C. D. (2014). Controlling the stream of thought: Working memory capacity predicts adjustment of mind-wandering to situational demands. *Psychonomic Bulletin & Review*, *21*(5), 1309-1315. https://doi.org/10.3758/s13423-013-0580-3

Rummel, J., & Nied, L. (2017). Do drives drive the train of thought?—Effects of hunger and sexual arousal on mind-wandering behavior. *Consciousness and Cognition*, *55*, 179-187. https://doi.org/10.1016/j.concog.2017.08.013

Rummel, J., Smeekens, B. A., & Kane, M. J. (2017). Dealing with prospective memory demands while performing an ongoing task: Shared processing, increased on-task focus, or both? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(7), 1047-1062. https://doi.org/10.1037/xlm0000359

Sayette, M. A., Reichle, E. D., & Schooler, J. W. (2009). Lost in the sauce: The effects of alcohol on mind wandering. *Psychological Science*, *20*(6), 747-752. https://doi.org/10.1111/j.1467-9280.2009.02351.x

Schad, D. J., Nuthmann, A., & Engbert, R. (2012). Your mind wanders weakly, your mind wanders deeply: Objective measures reveal mindless reading at different levels. *Cognition*, *125*(2), 179-194. https://doi.org/10.1016/j.cognition.2012.07.004

Schooler, J., & Schreiber, C. A. (2004). Experience, meta-consciousness, and the paradox of introspection. *Journal of Consciousness Studies*, *11*(7-8), 17-39. https://www.ingentaconnect.com/content/imp/jcs/2004/00000011/f0020007/art00003

Schooler, J. W. (2002). Re-representing consciousness: Dissociations between experience and meta-consciousness. *Trends in Cognitive Sciences*, *6*(8), 339-344. https://doi.org/10.1016/S1364-6613(02)01949-6

Schooler, J. W., Reichle, E. D., & Halpern, D. V. (2004). Zoning out during reading: Evidence for dissociations between experience and metaconsciousness. In D. T. Levin (Ed.), *Thinking and seeing: Visual metacognition in adults and children* (pp. 204–226). MIT Press.

Schooler, J. W., Smallwood, J., Christoff, K., Handy, T. C., Reichle, E. D., & Sayette, M. A. (2011). Meta-awareness, perceptual decoupling and the wandering mind. *Trends in Cognitive Sciences*, *15*(7), 319-326. https://doi.org/10.1016/j.tics.2011.05.006

Schubert, A. L., Frischkorn, G. T., & Rummel, J. (2019). The validity of the online thought-probing procedure of mind wandering is not threatened by variations of probe rate and probe framing. *Psychological Research*. Advance online publication. https://doi.org/10.1007/s00426-019-01194-2

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

Seli, P., Carriere, J. S. A., & Smilek, D. (2015). Not all mind wandering is created equal: Dissociating deliberate from spontaneous mind wandering. *Psychological Research*, *79*(5), 750-758. https://doi.org/10.1007/s00426-014-0617-x

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., Kane, M. J., Smallwood, J., Schacter, D. L., Maillet, D., Schooler, J. W., & Smilek, D. (2018). Mind-wandering as a natural kind: A family-resemblances view. *Trends in Cognitive Sciences*, *22*(6), 479-490. https://doi.org/10.1016/j.tics.2018.03.010

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

Seli, P., Smallwood, J., Cheyne, J. A., & Smilek, D.L. (2015). On the relation of mind wandering and ADHD symptomatology. *Psychonomic Bulletin & Review, 22*(3), 629-636. https://doi.org/10.3758/s13423-014-0793-0

Seli, P., Wammes, J. D., Risko, E. F., & Smilek, D. (2016). On the relation between motivation and retention in educational contexts: The role of intentional and unintentional mind wandering. *Psychonomic Bulletin & Review*, *23*(4), 1280-1287. https://doi.org/10.3758/s13423-015-0979-0

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron*, *79*(2), 217-240. https://doi.org/10.1016/j.neuron.2013.07.007

Shenhav, A., Cohen, J. D., & Botvinick, M. M. (2016). Dorsal anterior cingulate cortex and the value of control. *Nature Neuroscience*, *19*(10), 1286-1291. https://doi.org/10.1038/nn.4384

Shepherd, J. (2019). Why does the mind wander? *Neuroscience of Consciousness*, *2019*(1), Article niz014. https://doi.org/10.1093/nc/niz014

Singer, J. L. (1966). *Daydreaming: An introduction to the experimental study of inner experience*. Crown Publishing Group.

Singer, J. L. (1974). Daydreaming and the stream of thought: Daydreams have usually been associated with idleness and inattentiveness. Now, however, through an empirical research program, their general function and adaptive possibilities are being elucidated. *American Scientist*, *62*(4), 417-425. http://www.jstor.org/stable/27844990

Singer, J. L. (1975). Navigating the stream of consciousness: Research in daydreaming and related inner experience. *American Psychologist*, *30*(7), 727-738. https://doi.org/10.1037/h0076928

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*(1), 187-209. https://doi.org/10.2466/pms.1963.17.1.187

Singer, J. L., & Antrobus, J. S. (1970). *Manual for the Imaginal Processes Inventory*. Educational Testing Service.

Smallwood, J., & Andrews-Hanna, J. (2013). Not all minds that wander are lost: The importance of a balanced perspective on the mind-wandering state. *Frontiers in Psychology*, *4*, Article 441. https://doi.org/10.3389/fpsyg.2013.00441

Smallwood, J., Beach, E., Schooler, J. W., & Handy, T. C. (2008). Going AWOL in the brain: Mind wandering reduces cortical analysis of external events. *Journal of Cognitive Neuroscience*, *20*(3), 458-469. https://doi.org/10.1162/jocn.2008.20037

Smallwood, J., Brown, K. S., Tipper, C., Giesbrecht, B., Franklin, M. S., Mrazek, M. D., Carlson, J. M., & Schooler, J. W. (2011). Pupillometric evidence for the decoupling of attention from perceptual input during offline thought. *PLoS ONE*, *6*(3), Article e18298. https://doi.org/10.1371/journal.pone.0018298

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

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

Smallwood, J., McSpadden, M., & Schooler, J. W. (2008). When attention matters: The curious incident of the wandering mind. *Memory & Cognition*, *36*(6), 1144-1150. https://doi.org/10.3758/mc.36.6.1144

Smallwood, J., Nind, L., & O'Connor, R. C. (2009). When is your head at? An exploration of the factors associated with the temporal focus of the wandering mind. *Consciousness and Cognition*, *18*(1), 118-125. https://doi.org/10.1016/j.concog.2008.11.004

Smallwood, J., & O'Connor, R. C. (2011). Imprisoned by the past: Unhappy moods lead to a retrospective bias to mind wandering. *Cognition and Emotion*, *25*(8), 1481-1490. https://doi.org/10.1080/02699931.2010.545263

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

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

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

Smilek, D., Carriere, J. S. A., & Cheyne, J. A. (2010). Out of mind, out of sight: Eye blinking as indicator and embodiment of mind wandering. *Psychological Science*, *21*(6), 786-789. https://doi.org/10.1177/0956797610368063

Son, L. K., & Schwartz, B. L. (2002). The relation between metacognitive monitoring and control. In T. J. Perfect & B. L. Schwartz (Eds.), *Applied metacognition* (pp. 15–38). Cambridge University Press. https://doi.org/10.1017/CBO9780511489976.003

Stawarczyk, D., Cassol, H., & D'Argembeau, A. (2013). Phenomenology of future-oriented mind-wandering episodes. *Frontiers in Psychology*, *4*, Article 425. https://doi.org/10.3389/fpsyg.2013.00425

Stawarczyk, D., Majerus, S., & D'Argembeau, A. (2013). Concern-induced negative affect is associated with the occurrence and content of mind-wandering. *Consciousness and Cognition*, *22*(2), 442-448. https://doi.org/10.1016/j.concog.2013.01.012

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

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

Strick, M., Dijksterhuis, A., Bos, M. W., Sjoerdsma, A., van Baaren, R. B., & Nordgren, L. F. (2011). A meta-analysis on unconscious thought effects. *Social Cognition*, *29*(6), 738-762. https://doi.org/10.1521/soco.2011.29.6.738

Taruffi, L., Pehrs, C., Skouras, S., & Koelsch, S. (2017). Effects of sad and happy music on mind-wandering and the default mode network. *Scientific Reports*, *7*, Article 14396. https://doi.org/10.1038/s41598-017-14849-0

Teasdale, J. D., Segal, Z., & Williams, J. M. G. (1995). How does cognitive therapy prevent depressive relapse and why should attentional control (mindfulness) training help? *Behaviour Research and Therapy*, *33*(1), 25-39. https://doi.org/10.1016/0005-7967(94)E0011-7

Unsworth, N., & McMillan, B. D. (2013). Mind wandering and reading comprehension: Examining the roles of working memory capacity, interest, motivation, and topic experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(3), 832-842. https://doi.org/0.1037/a0029669

Unsworth, N., & Robison, M. K. (2016). Pupillary correlates of lapses of sustained attention. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(4), 601-615. https://doi.org/10.3758/s13415-016-0417-4

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

Vinski, M. T., & Watter, S. (2012). Priming honesty reduces subjective bias in self-report measures of mind wandering. *Consciousness and Cognition*, *21*(1), 451-455. https://doi.org/doi.org/10.1016/j.concog.2011.11.001

Wammes, J. D., & Smilek, D. (2017). Examining the influence of lecture format on degree of mind wandering. *Journal of Applied Research in Memory and Cognition*, *6*(2), 174-184. https://doi.org/10.1016/j.jarmac.2017.01.015

Waroquier, L., Marchiori, D., Klein, O., & Cleeremans, A. (2010). Is it better to think unconsciously or to trust your first impression? A reassessment of unconscious thought theory. *Social Psychological and Personality Science*, *1*(2), 111-118. https://doi.org/10.1177/1948550609356597

Watkins, E. R. (2008). Constructive and unconstructive repetitive thought. *Psychological Bulletin*, *134*(2), 163-206. https://doi.org/10.1037/0033-2909.134.2.163

Weinstein, Y. (2018). Mind-wandering, how do I measure thee with probes? Let me count the ways. *Behavior Research Methods*, *50*(2), 642-661. https://doi.org/10.3758/s13428-017-0891-9

Weinstein, Y., De Lima, H. J., & van der Zee, T. (2018). Are you mind-wandering, or is your mind on task? The effect of probe framing on mind-wandering reports. *Psychonomic Bulletin & Review*, *25*(2), 754-760. https://doi.org/10.3758/s13423-017-1322-8

Wiemers, E. A., & Redick, T. S. (2019). The influence of thought probes on performance: Does the mind wander more if you ask it? *Psychonomic Bulletin & Review*, *26*(1), 367-373. https://doi.org/10.3758/s13423-018-1529-3

Yang, S.-n. (2006). An oculomotor-based model of eye movements in reading: The competition/interaction model. *Cognitive Systems Research*, *7*(1), 56-69. https://doi.org/10.1016/j.cogsys.2005.07.005

Yanko, M. R., & Spalek, T. M. (2014). Driving with the wandering mind: The effect that mind-wandering has on driving performance. *Human Factors*, *56*(2), 260-269. https://doi.org/10.1177/0018720813495280

Zedelius, C. M., Broadway, J. M., & Schooler, J. W. (2015). Motivating meta-awareness of mind wandering: A way to catch the mind in flight? *Consciousness and Cognition*, *36*, 44-53. https://doi.org/10.1016/j.concog.2015.05.016

Zedelius, C. M., & Schooler, J. W. (2015). Mind wandering "Ahas" versus mindful reasoning: Alternative routes to creative solutions. *Frontiers in Psychology*, *6*, Article 834. https://doi.org/10.3389/fpsyg.2015.00834

Zedelius, C. M., & Schooler, J. W. (2017). What are people's lay theories about mind wandering and how do those beliefs affect them? In C. M. Zedelius, B. C. N. Müller, & J. W. Schooler (Eds.), *The science of lay theories: How beliefs shape our cognition, behavior, and health* (pp. 71-93). Springer International Publishing. https://doi.org/10.1007/978-3-319-57306-9_4

Zeigarnik, B. (1927). Über das Behalten von erledigten und unerledigten Handlungen. *Psychologische Forschung*, *9*, 1-85.

Zhang, H., Miller, K., Sun, X., & Cortina, K. S. (2020). Wandering eyes: Eye movements during mind-wandering in video lectures. *Applied Cognitive Psychology*, *34*(2), 449-464. https://doi.org/10.1002/acp.3632

Zhiyan, T., & Singer, J. L. (1997). Daydreaming styles, emotionality and the Big Five personality dimensions. *Imagination, Cognition and Personality*, *16*(4), 399-414. https://doi.org/10.2190/ateh-96ev-exyx-2adb

**List of Figures**

**Appendix A1 – Manuscript 1**

Note:

This is a previous version of an article that has been published in its final version in *Behavior Research Methods* [© Springer Nature].

Steindorf, L., & Rummel, J. (2020). Do your eyes give you away? A validation study of eye-movement measures used as indicators for mindless reading. *Behavior Research Methods, 52*(1), 162-176. https://doi.org/10.3758/s13428-019-01214-4

# Do your eyes give you away? A validation study of eye-movement measures used as indicators for mindless reading

Lena Steindorf & Jan Rummel

Heidelberg University

Authors' Note

Lena Steindorf & Jan Rummel, Department of Psychology, Heidelberg University, Heidelberg, Germany. The present research was supported by a grant from the German research foundation DFG to the second author (grant no: RU1996/1-1). The authors would like to thank Nadine Gronewold, Amelie Haindl, Fynn Ole Wöstenfeld, Fabian Dittmar, and Holly Hammerton for their help in data collection with special thanks to Holly for also proofreading the manuscript.

Correspondence concerning this article should be addressed to Lena Steindorf, Department of Psychology, Heidelberg University, Hauptstrasse 47-51, 69117 Heidelberg, Germany. E-mail: lena.steindorf@psychologie.uni-heidelberg.de

**Abstract**

Identifying eye-movement measures as objective indicators of mind wandering seems to be a work in progress. We reviewed research comparing eye movements during self-categorized episodes of normal versus mindless reading and found little consensus regarding the specific measures that are sensitive to attentional decoupling during mind wandering. To address this issue of inconsistency, we conducted a new, high-powered eye-tracking experiment and considered all previously identified mind-wandering indicators. In our experiment, only three measures (reading time, fixation count, and first-fixation duration) positively predicted self-categorized mindless reading. Asides from these single measures, the word-frequency effect was found to be generally less pronounced during mindless-reading than during normal-reading episodes. To additionally test for convergent validity between objective and subjective mind-wandering measures, we utilized eye-movement measures as well as thought reports to examine the effect of meta-cognitive awareness on mind-wandering behavior. We expected participants anticipating a difficult comprehension test to mind wander less during reading compared to those anticipating an easy test. Although we were able to induce meta-cognitive expectancies about task difficulty, there was no evidence that difficulty expectancies affected either subjectively-reported or objectively-measured mind-wandering.

*Keywords*: mind wandering, task-unrelated thought, eye movements, reading

Do your eyes give you away? A validation study of eye-movement measures used as indicators for mindless reading

Imagine you are reading a book. Your eyes are moving across the page, line by line. While processing words and sentences, you immerse yourself deeper and deeper into the story. You have been reading for a while now, and your eyes keep on scanning the pages, from top to bottom, from left to right. Suddenly you realize you have just been thinking about that huge fight you had with your best friend. It seems impossible for you to reproduce the last text passages, although you remember your eyes moving across the lines. You were mind wandering. That is, your thoughts trailed off from the task at hand, which was, in this case, reading for comprehension. When people mind wander, they think about personal problems, unfulfilled tasks, or other things unrelated to their current task (Schooler et al., 2011). Mind-wandering episodes often occur without intention or even awareness (Smallwood & Schooler, 2006). For this reason, it is possible for you to mindlessly "read" a considerable amount of lines without noticing that your thoughts are trailing off. The fact that you maintain a reading-like eye-movement behavior while doing so makes mindless reading an attractive research topic in the mind-wandering domain. What happens to eye movements when attention is directed away from ongoing text processing? Do eye movements differ systematically between normal and mindless reading in such a way that they could serve as objective, non-intrusive indicators of mind wandering? In the present work, we tried to answer these questions by reviewing research comparing eye movements during self-categorized episodes of normal versus mindless reading. We also report a newly conducted, high-powered eye-tracking experiment, with the objective of validating previously identified mind-wandering indicators.

Mind wandering is a ubiquitous phenomenon. As much as 30-50 % of all daily thoughts are unrelated to current external events (Kane et al., 2017; Killingsworth & Gilbert, 2010; Klinger, 1999). When thoughts trail off, the perception of the external world is affected.

Thoughts about a fight you had with your best friend move your mental focus away from the words and sentences on the page in front of you. During mind-wandering episodes, attention is directed away from external, perceptual input towards internal processing (Schooler et al., 2011; Smallwood & Schooler, 2006; Stawarczyk, Majerus, Maquet, & D'Argembeau, 2011). When your thoughts trail off during reading, you no longer pay sufficient attention to the lines of text. This attentional decoupling process does not seem to work in an all-or-nothing fashion. Schad, Nuthmann, and Engbert (2012) proposed the *levels-of-inattention hypothesis*. Their findings support the idea of graded attentional decoupling at different levels of the cognitive hierarchy: Processing of external information can fail at early (*deep decoupling*, cascading down to later levels, see Smallwood, 2011) but also at late (*weak decoupling*) perceptual levels. We believe this extension of a dichotomous view of perceptual decoupling to be a very fruitful approach. For our purpose, however, we took a step back and critically reviewed research comparing eye movements during text-focused versus decoupled attention (mind wandering), regardless of decoupling levels.

An attention shift towards internal processing renders external information-encoding errors more likely (Smallwood, Baracaia, Lowe, & Obonsawin, 2003), thus causing performance decrements in external tasks (Christoff, Gordon, Smallwood, Smith, & Schooler, 2009). During reading, mind wandering is assumed to be reflected by attentional decoupling from the semantic input, meaning the lines of text one is reading. Consequently, reading comprehension is very likely to suffer. Previous research has confirmed that thoughts often trail off during reading and that high mind-wandering frequencies are associated with poorer reading comprehension (Schooler, Reichle, & Halpern, 2004). It is assumed that readers fail to mentally connect events, or linguistic units, to create a situational model of the story during task-unrelated-thought (TUT) episodes (Smallwood, McSpadden, & Schooler, 2008).

But what does attentional decoupling have to do with eye movements, and how may eye movements aid us in detecting mindless reading? The basic idea behind eye-tracking research in general is that what people are looking at reflects what they are mentally processing, at least to some degree (Uzzaman & Joordens, 2011). During reading—a behavior that can be broken down into a sequence of fixations and saccades—a fixation on a given word is assumed to reflect the mental processing of this word. This logic becomes apparent when taking a look at the robust word-frequency effect (Inhoff & Rayner, 1986). Identifying and processing a very uncommon word such as “vat” takes longer than processing a very common word such as “cat.” Fixation times are reflective of the need for more intense processing: Low-frequency words are fixated longer than high-frequency words. Cognitive-control theories (Just & Carpenter, 1980), in contrast to oculomotor-control theories (Yang, 2006), posit a close *eye-mind link*. Models such as SWIFT (Engbert, Nuthmann, Richter, & Kliegl, 2005) and E-Z Reader (Pollatsek, Reichle, & Rayner, 2006) assume that cognitive and oculomotor processes are linked to *cognitive forces* driving eye-movement behavior during normal reading. Therefore, cognitive-control theories would predict differences in eye movements during normal versus mindless reading. The eye-mind link should be interrupted when the mind starts to wander away from external text processing towards internal processing. During mind wandering, cognitive forces are engaged in thoughts unrelated to the to-be-read text, limiting their power to drive eye movements. Mindless readers may thus move their eyes beyond the word “vat” without thoroughly processing it—and thus not taking the additional fixation time its processing usually requires. That is, when attention shifts from lines of text towards inner thoughts, eye movements should be (partially) decoupled from ongoing text processing, resulting in less sensitivity towards lexical features, as for example word frequency. Research applying the *z*-string reading paradigm (Nuthmann & Engbert, 2009; Rayner & Fischer, 1996), during which participants mindlessly read “words” only containing the letter *Z* (e.g., “Zzzzz zz zzzzzz”), suggests that readers are not only less sensitive to lexical variables, but that single eye-movement measures, such as fixation duration, are erratic (in this case, longer) when thoughts are off-task. Furthermore, Smilek, Carriere, and Cheyne (2010) noted that attentional decoupling during mindless reading could also be represented by a higher blinking frequency, thereby even physically reducing the processing of external input.

This reasoning motivated several studies (see below) trying to identify eye-movement measures that are sensitive to attentional decoupling during mindless reading. Because people are not able to consciously control or manipulate fixations and saccades (Rayner, 1998), mind-wandering research would greatly benefit if eye-movement measures were found to be robust and stable indicators of TUT episodes. Eye movements could be used as biological markers in addition to or as a replacement for subjective self-reports. Until now, mind wandering has typically been measured via self-reports. In many mind-wandering experiments, participants are asked to briefly describe and/or classify their current thoughts' content when a probe interrupts their present task. Although these self-reports proved to be reasonably valid (e.g., McVay & Kane, 2012), they might still be prone to memory-driven and/ or classification errors. Also, the demand characteristics of a given task might influence subjectively reported mind wandering (Vinski & Watter, 2012). The apparent subjectivity of self-reports has thus led to a search for objective biological or behavioral indicators of TUTs. If found, these could not only be used to further validate subjective thought reports, but also to directly measure mind-wandering behavior without a subjective component. Besides eye movements, reaction times seem to be promising objective markers of TUTs. McVay and Kane (2012) measured attentional lapses using thought probes and found that these lapses go along with higher $\tau$ parameters, which represent the tail of reaction time distributions. In other words, more mind wandering goes along with higher proportions of extremely long reaction times. Of course, a researcher's choice of objective markers is task-dependent. For go/no-go tasks like the SART (Robertson, Manly, Andrade, Baddeley, & Yiend, 1997), which is often used by mind-wandering researchers (McVay & Kane, 2012; Unsworth & McMillan, 2014), reaction times could be the appropriate means and seem to be a fruitful approach. For reading tasks, eye movements have started to emerge as biological mind-wandering markers. In the following sections, we review five studies that aimed to identify eye-movement measures as objective mind-wandering indicators. To this end, we will give an overview on the methods that have been applied in this line of research and discuss the results that were obtained (for an overview, see Table 1).

For the present purpose, we only review studies in which participants' thoughts were randomly probed (probe-caught method, see Schooler et al., 2004) during reading. The probe-caught method does not rely on participants' awareness of their mind-wandering behavior and is supposed to be a reliable estimate of mind-wandering frequency (Smallwood & Schooler, 2006). In the to-be-reviewed studies, self-categorized normal-reading episodes were compared to self-categorized mindless-reading episodes in order to determine changes in eye-movement behavior. This approach is especially useful in order to make eye-movement measures utilizable as objective mind-wandering indicators, because it allows the validation of possible mind-wandering indicators by means of the most widely used mind-wandering assessment methods, namely online thought probes.

Reichle, Reineberg, and Schooler (2010) had four participants read the entirety of *Sense and Sensibility* by Jane Austen in up to 15 hour-long sessions, while tracking their eye movements and applying the probe-caught method to assess mind-wandering behavior, with a total of 151.5 probes per participant, on average. The authors additionally applied the self-caught method, which requires participants to press a certain key whenever they catch themselves mind wandering[1]. Eye-movement measures for six different time intervals

---

[1] Although the results differed for self- versus probe-caught mind-wandering, we focus solely on probe-caught results in the present work. Distinguishing between self- and probe-caught mind-wandering provides important insights into awareness processes and further research will be needed to determine the influence of different levels of mind-wandering awareness on eye-movement behavior (see also Schad, Nuthmann and Engbert, 2012). Still, in most of the studies we reviewed, the probe-caught method was applied exclusively. Thus, we will focus on this method alone for comparability reasons.

preceding the probes (2.5, 5, 10, 30, 60, and 120s prior to a thought probe) were analyzed regarding differences between episodes of self-classified normal versus mindless reading. In a similar study, Smilek et al. (2010) analyzed data from 12 participants who had read two passages of *A Short History of Nearly Everything* by Bill Bryson with their eye movements being tracked and their thoughts being probed 20 times. The 5 s time intervals preceding thought probes were analyzed. Uzzaman and Joordens (2011) also analyzed eye-movement measures from the 5 s time intervals preceding thought probes. Data from 22 participants who had read 16 pages of *War and Peace* by Tolstoy entered their analysis. Again, the probe-caught method was applied, and participants received a total of ten probes per participant on average. A slightly different method to detect differences in eye-movement behavior between normal and mindless reading was employed by Foulsham, Farley, and Kingstone (2013). In their experiment, not specific time intervals, but single sentences were analyzed. 26 participants read 120 sentences, 48 of which being to-be-examined key sentences. Thoughts were probed approximately nine times with a probe always appearing after a key sentence. In the aforementioned three experiments, probes appeared randomly during full-text reading, resulting in different to-be-analyzed text passages for each participant and thus increasing error variance. The single-sentence-reading method by Foulsham et al. kept the to-be-examined reading materials equal for all participants and even allowed for the manipulation of the linguistic input prior to thought probes. Participants read sentences containing either high- or low-frequency words, so that the influence of mind wandering on the word-frequency effect could be examined, which turned out to be smaller during mindless reading in this study. A potential disadvantage of this method is, however, that presenting single sentences is less ecologically valid than full-text reading. In a study by Frank, Nara, Zavagnin, Touron, and Kane (2015), participants again read chapters from Tolstoy's *War and Peace*. 29 younger adults'[2] eye-tracking data from 3-8s time intervals preceding 20 thought probes were analyzed.

By and large, the methods used in these five studies show a reasonable degree of similarity. The probe-caught method was always applied to identify passages read normally versus mindlessly, and eye-tracking data were compared for these passages. Methodologically, these studies differed concerning the lexical input (full text versus single sentences, different book chapters, number of lines on each screen, etc.), the total reading time (from less than an hour up to 15 hours), the probing procedure (frequency of probing, additionally assessed self-caught mind wandering, number and selection of response options, etc.), the number of participants (four to 29), and the analyzed areas of interest preceding thought probes (a single sentence versus a specific time interval). Some of these specific features might have had an influence on the assessment of mind wandering and the eye-movement behavior. For example, differences in reported mind wandering can arise due to the framing and wording of thought probes and response options (Weinstein, 2018; Weinstein, De Lima, & van der Zee, 2018). Additionally, eye-movement behavior might have been influenced by the length of the experiment (e.g., due to fatigue) or by the lexical input (e.g., single sentences versus full text).

At least for studies that employ similar methods, we would expect to find converging eye-movement patterns for normal versus mindless reading if eye movements were robust indicators of mind wandering. Interestingly, all five studies varied widely regarding the selected eye-movement measures and the obtained results. Table 1 shows eye-movement measures that were found to differ significantly ($p < .05$) between normal and mindless-reading episodes in at least one of the previously described studies. For comparability reasons, the table is limited to effects found for younger adults with the probe-caught method

[2] The authors tested older and younger adults. For comparability reasons, we only report the effects found for younger adults.

assessing TUT episodes, excluding task-related inferences that were additionally assessed only in the Frank et al. (2015) study. Whenever a significant effect was reported for a measure in any study, we checked all other studies as to whether they reported results for this measure. As becomes apparent from Table 1, there seem to be large differences between studies: Significant effects were found for several measures, but studies strongly differ as to which eye-movement measures were selected to be analyzed. Certainly, this is partially due to the theoretical focus of the respective study. However, to make eye-movement measures utilizable as objective mind-wandering indicators, a greater degree of consensus in the selection of eye-movement measures of interest is desirable. Importantly, even if the same measure was analyzed in two or more studies, effects could not always be replicated and even showed opposite directions. For example, Smilek et al. (2010) found a lower fixation count during mindless reading compared to normal reading, whereas Foulsham et al. (2013) found a higher fixation count during mindless reading.

Such inconsistencies in results can indicate validity problems for eye-movement measures as indicators for mindless reading. However, they can also reflect meaningful processing differences due to methodological specifics. Therefore, discrepancies in results between studies do not have to be a knockout argument against the usefulness of eye-movements as mind-wandering indicators. For example, single-sentence reading like in the Foulsham et al. (2013) study likely imposes different cognitive demands as full-text reading. Although different contextual demands might generate discrepancies in results between studies, objective indicators should converge when contexts are similar. When contexts differ, boundary conditions should be identified, so that it becomes possible to predict when and why specific patterns of results will occur. To this end, studies manipulating the contextual demands of interest (e.g., single-sentence vs. full-text reading) are required.

Notably, even for studies that applied similar methods and whose contextual demands appeared to be relatively comparable, Table 1 lists discrepancies in results between studies.

To address these inconsistencies, we conducted a new, high-powered eye-tracking experiment, in which we confined the analyses to the eye-movement measures that were previously identified as potential mind-wandering indicators (i.e., those listed in Table 1) a priori. In this study, we applied the standard approach (see above) and asked participants to read two chapters of Oliver Sacks's *Musicophilia: Tales of Music and the Brain* (Sacks, 2008) for comprehension, and further asked them to respond to ten thought probes while reading. We tracked participants' eyes during reading so that we could compare eye-movement behavior between normal and mindless reading.

Besides the validation of eye-movement measures as mind-wandering markers, the present study's secondary aim was to actually apply these markers to examine the effect of meta-cognitive expectancies on mind wandering. Previous research has shown that mind-wandering behavior during cognitive tasks changes in accordance with varying task demands. That is, mind wandering is usually reduced when task demands increase (e.g., Rummel & Boywitt, 2014). Considering that mind wandering is beneficial in some situations, this adjustment seems to be adaptive in everyday life. TUTs are often future related and focus on unfulfilled tasks or personal problems. In this way, they might help people plan future actions (Klinger, 1999; Mooneyham & Schooler, 2013), come up with creative ideas (Baird et al., 2012; but see Smeekens & Kane, 2016), or keep future task goals active (Steindorf & Rummel, 2017). Given these potential positive effects, and considering the *context-regulation hypothesis* (Smallwood & Andrews-Hanna, 2013), efficient adjustment should allow people to benefit from TUTs as long as the situation allows or encourages it, while also minimizing the costs of mind wandering when working on demanding tasks.

However, directly manipulating the task demands for a reading task would have required us to change the to-be-read text between experimental conditions. Instead, we decided to focus on meta-cognitive awareness, the knowledge about one's own cognitive processes (Flavell, 1979), as one possible underlying factor of mind-wandering adjustment. We manipulated participants' expectancies regarding the difficulty of an upcoming reading comprehension test. That is, all participants read the same text but some expected a difficult reading comprehension test afterwards, whereas others expected an easy test. During the reading phase, we tracked participants' eye movements and periodically probed their thoughts. All participants then worked on the same reading comprehension test. Analogous to the results found for high task demands (e.g., Rummel & Boywitt, 2014), we expected participants with high-difficulty expectations to mind wander less than low-difficulty-expectation participants. Convergent evidence for such an effect from subjective thought-reports and (some) objective eye-movement measures (see above and Table 1) would further speak for the convergent validity of the respective measures. In addition, we expected participants with high-difficulty expectations to perform better in the reading comprehension test than those with low-difficulty expectations.

Across the subsequent *Method* and *Results* sections, we report how we determined our sample size and all data exclusions, manipulations, and measures in the study (Simmons, Nelson, & Simonsohn, 2012). The study reported in this article was preregistered on June 28, 2016 (https://osf.io/sbeqn/) and, if not otherwise indicated, we followed the research and analysis protocol as stated in the preregistration.

## Method

### Participants, data cleansing, and design

For multilevel-modeling approaches, there is currently no common understanding regarding power analyses. However, simulations by Maas and Hox (2005) determined what constitutes a sufficient sample size for accurate parameter estimation. The authors recommend, as stated in the preregistration, collecting data of at least 50 second-level and seven first-level units. Considering the accuracy benefits (for variance estimation) that an increasing number of second-level units might offer (beyond 50, see also Paccagnella, 2011), but also the effort of an eye-tracking study, 122 participants (second-level units) were tested at Heidelberg University, Germany. Using ten measurement points per participant, we ensured that our design also met the criteria of at least seven first-level units. To account for potential participant exclusions due to eye-tracking related problems, we tested 22 participants more than we had preregistered.

Eye-tracking data of one participant were lost due to a software problem. The eye-tracking data quality for all other participants was determined sentence by sentence for all participants and all target sentences by two independent raters. A sentence's eye-tracking data were categorized as unusable when technical errors were clearly evident (e.g., when calibration had failed), when a strong drift at the beginning or end of the lines was visible, or when a participant's eyes accidently hit a thought probe trigger (see below) before reading the target sentence causing a premature thought probe appearance. For 88.96 % of all target sentences, both raters agreed on the categorization. For the other target sentences, a liberal decision was made, and the eye-tracking data were categorized as useable. Fourteen participants were excluded from the final data set due to their producing predominantly unusable eye-tracking data (less than five useable target sentences), resulting in $N = 107$ ($M_{age} = 22.58$, $SD_{age} = 4.01$, 78.50 % female, normal or corrected-to-normal vision) for all further analyses. For the final data set, 93.77 % of all target sentences' eye-tracking data were categorized as usable. Unusable data were excluded from further analyses, relying on multilevel models' beneficial characteristic of being applicable to data sets with randomly missing values.

We used a one-factorial design to investigate the influence of difficulty expectations (expectation of an easy versus a difficult comprehension test) on mind-wandering behavior and reading performance.

## Apparatus

Eye movements were recorded using the SMI RED500 system (SensoMotoric Instruments, Potsdam, Germany) with a sampling rate of 500 Hz and a spatial resolution

(RMS) of 0.03°. Using a remote eye-tracking system, free head movements (in a 40 cm × 20 cm range) allowed for naturalistic reading, but participants were asked to move around as little as possible in 60 cm to 80 cm distance from the monitor (22″). The software BeGaze's (SensoMotoric Instruments, Potsdam, Germany) algorithm was used to calculate eye movement measures, using a velocity threshold of 40 °/s to identify saccades[3].

**Materials**

**Reading task and target sentences.** Participants read a shortened German version of two chapters of Oliver Sacks's *Musicophilia: Tales of Music and the Brain* (Sacks, 2008). The text was presented on 58 pages in black-on-gray 32-point Arial font with double spacing and a maximum of seven lines per page, using the software Experiment Center (SensoMotoric Instruments, Potsdam, Germany). Participants moved through the text autonomously using the space bar to turn pages. On each page, a participant's fixation on a trigger cross (dwell time = 500 ms) in the top left corner made the text appear. This procedure ensured that, for each page, eye-movement recordings always started at the top left corner position. Ten target sentences (always presented in the middle of the page, $M_{\text{word-count}} = 31.30$; $SD_{\text{word-count}} = 5.27$) were determined before data collection, five containing a high-frequency ($M_{\text{frequency}} = 1892.80$, $SD_{\text{frequency}} = 1309.78$), the other five containing a low-frequency ($M_{\text{frequency}} = 8.40$, $SD_{\text{frequency}} = 10.90$) target word. Word frequencies were extracted from the *dlex* database (Heister et al., 2011) and target words were matched for length (see Appendix for a full list of target words including translations and information on frequency and length). After reading a target sentence, participant's fixations on the first words of the following sentence (dwell time

[3] From the eye-tracking and mind wandering studies we reviewed, only one (Foulsham et al., 2013) provided details about the event detection method that was applied. Because these authors used a velocity-based algorithm and none of the other studies' authors indicated that they deviated from this standard procedure, we also employed the standard algorithm as well as parameter values our software recommends for comparability reasons. We acknowledge, however, that more sophisticated methods for event detection are currently under development (e.g. Nyström & Holmqvist, 2010) that will eventually replace this method in the near future.

= 300 ms) triggered the appearance of a thought probe. After responding to the thought probe, participants were redirected to the page they had been reading before the probe appeared. In doing so, we were able to define target sentences with specific features (high- versus low-frequency words) and to keep to-be-examined passages equal for all participants (Foulsham et al., 2013). Still, participants could engage in natural reading in contrast to the Foulsham et al. study in which single sentences were presented sequentially.

**Thought probes.** To subjectively assess mind wandering, participants' thoughts were probed after each target sentence resulting in a total of ten probes. Participants were asked to categorize their current thoughts as being task-related thoughts ("I am thinking about the text I am reading," later referred to as TRTs), task-triggered intrusions ("I am thinking about things related to the text I am reading," later referred to as TTIs), or TUTs ("I am thinking about things unrelated to the text").

**Procedure**

In individual sessions, participants signed a consent form before receiving detailed written information about the eye-tracking procedure and instructions for the reading phase and the comprehension test. Furthermore, the concept of mind wandering, and the response options of the thought probes were explained. A five-point-calibration of the eye tracker was applied prior to the reading phase and was repeated once in the middle of the reading phase, as well as whenever considered necessary by the experimenter who monitored eye movements on a separate screen, nonvisible to the participants. Once the eye tracker was successfully calibrated, participants practiced turning pages in a self-paced fashion on a passage of instruction text, by pressing the space bar and fixating the trigger area on the following page. They also practiced the thought-probing procedure once before receiving critical instructions for the experimental manipulation. Participants in the high-difficulty-expectation condition were told that the upcoming comprehension test was going to be very difficult for a student population whereas participants in the low-difficulty-expectation condition were informed that the test was going to be fairly easy. This information was embedded in a cover-story referring to previous research to render the manipulation more trustworthy. Following these instructions, the reading phase began and participants autonomously read the two chapters within up to 30 minutes, and responded to a total of ten thought probes, with each probe occurring after a target sentence. After the reading task, participants completed a retrospective mind-wandering questionnaire categorizing the entirety of their thoughts during the reading phase into the three thought probe response categories (TRTs, TTIs, TUTs) using percentage scores. We then asked participants to rate how difficult they expected the upcoming reading comprehension test to be on a ten-point scale, from *very easy* to *very difficult*. The following test consisted of 15 multiple-choice items (four response options each, only one of which was correct) which had been experienced as moderately difficult in a pretest. After completion of the comprehension test, participants rated how difficult they had perceived the test to be on the previously used ten-point scale and answered demographic questions. Finally, participants were debriefed and dismissed.

## Results

The analyses reported in this section were executed as preregistered, if not indicated otherwise. We employed confirmatory strategies of analyses and conducted one-tailed tests (Cho & Abe, 2013) whenever we had preregistered a directional hypothesis, that is for all analyses concerning the difficulty-expectation manipulation. For these analyses, we considered any effect in the direction opposite from our preregistered expected outcomes to be non-significant. For those analyses concerning the validation of eye-movement measures (e.g., because previous research had found opposing results), we employed a standard non-directional hypothesis-testing strategy.

### Behavioral measures

Performance on the reading comprehension test was calculated as the percentage of correctly answered comprehension questions. The amount of online-reported TUTs and TTIs during the reading phase was defined as the sum of probes in which participants categorized their thoughts as being either task-unrelated or task-triggered, respectively.

Table 2 shows means, standard deviations, and correlations for the behavioral measures. Mean comprehension accuracy was satisfactory, implying that participants generally paid a decent amount of attention to the text, although, on average, TUTs or TTIs were reported in almost half of all thought-probe occurrences. Retrospective thought reports converged with online thought reports, as indicated by the moderate to strong positive correlations of the two. Comprehension accuracy weakly correlated with TUTs (online and retrospective); the better performance in the comprehension test was, the fewer TUTs were reported. The comprehension test's expected difficulty correlated weakly with retrospectively reported TUTs: The fewer TUTs were reported, the more difficult the test was expected to be. A closer look at this pattern revealed that it was only present for participants in the high-difficulty-expectation group. For this group, difficulty expectations negatively correlated with online TUTs, $r(52) = -.31$, $p = .011$ (one-tailed), and, numerically identically, with retrospectively reported TUTs, $r(52) = -.31$, $p = .011$ (one-tailed). For the low-difficulty-expectation group, we found close-to-zero correlations.

To test for further group differences on behavioral measures, we first ran a manipulation check. A 2 × 2 analysis of variance (ANOVA) with experimental condition (high- versus low-difficulty-expectation) as between-participants and point in time (before versus after test) as within-participants factor for comprehension test difficulty estimates revealed a significant difference between expected (before test, $M = 6.42$, $SD = 2.01$) and perceived (after test, $M = 5.46$, $SD = 1.71$) difficulty, $F(1, 105) = 16.68$, $p < .001$, $\eta^2_p = .14$. The main effect of experimental condition was also significant, $F(1, 105) = 33.22$, $p < .001$, $\eta^2_p = .24$, with higher difficulty estimates for high-difficulty-expectation participants ($M = 6.57$, $SD = 1.12$) than for low-difficulty-expectation participants ($M = 5.30$, $SD = 1.17$). A significant interaction, $F(1, 105) = 72.49$, $p < .001$, $\eta^2_p = .19$, and additional simple effects revealed that a group difference was only present for expected difficulty estimates, $F(1, 105) = 62.27$, $p < .001$, $\eta^2_p = .37$, and not for perceived difficulty estimates, $F(1, 105) < 1$, $p = .734$, $\eta^2_p < .01$. High-difficulty-expectation participants ($M = 7.63$, $SD = 1.41$) expected the comprehension test to be more difficult than low-difficulty-expectation participants ($M = 5.19$, $SD = 1.77$), indicating a successful expectation manipulation. However, contradictory to our predictions, we did not find any significant group differences for online TUTs, online TTIs, retrospective TUTs, retrospective TTIs, or comprehension accuracy, all $t$s $\leq 1.1$, all $p$s $>.270$. Taken together, although participants in the high-difficulty-expectations condition expected the comprehension test to be more difficult than those in the low-difficulty-expectation condition, we did not find corresponding differences regarding either their mind-wandering behavior during the reading phase or their test performance. Therefore, this manipulation was not successful in influencing mind-wandering behavior in any way.[4] On that account, we refrained from testing the preregistered mediating effect of mind wandering on reading comprehension.

**Eye-tracking analyses**

**Comparison of eye-movement measures during normal reading versus TUT episodes**

We used the software BeGaze's (SensoMotoric Instruments, Potsdam, Germany) algorithm to calculate all eye-movement measures of interest (Table 3) from the raw eye-tracking data. For each eye-tracking measure, up to ten data points (depending on data exclusions) per participant were considered for the following analyses. For within-word-regression count, between-word-regression count, blink count, and fixation count, the number of observed occurrences of the respective event (i.e., regressions, blinks, or fixations) during a given target sentence and for a given participant represents one data point. For reading time,

[4] An additional multilevel regression analysis for TUT occurrences with the predictors condition (effect-coded), difficulty expectation (as continuous predictor), and their interaction also showed no effect of difficulty expectancies on TUTs. We thank Jonathan Smallwood for suggesting this additional analysis.

the time it took a participant to read a given target sentence represents one data point. Last, for first-fixation duration, the mean of the entirety of a participant's first fixations in one target sentence represents a data point. That is, all values are to be interpreted sentence-wise (e.g., two blinks per sentence), except for first-fixation duration, which is to be interpreted word-wise (e.g. an average first-fixation duration of 150 ms for all words in one target sentence).

Table 3 shows means (aggregated across target sentences and participants) and standard deviations for all relevant unstandardized and uncorrected eye movement measures for TRT and TUT episodes, respectively. Before sentence-wise calculated eye-movement measures were entered in the analysis, they were divided by the respective target sentence's character count (to account for different sentence lengths) and z-standardized (to make the estimated coefficients comparable). The first-fixation-duration measure was also z-standardized.

To test for differences in eye-movement measures between TUT and TRT episodes, we used a multilevel modeling approach. We chose this approach for several reasons: First, it can account for dependencies in the data due to repeated measures. In our case, eye-movement measures and thought probe responses were nested in target sentences and participants. Second, because mind wandering is known to become more likely the more time is spent on a task (e.g., Foulsham et al., 2013; McVay & Kane, 2009; Rummel & Nied, 2017; Steindorf & Rummel, 2017), we also considered time-on-task as an additional predictor. Because, analogously to these previous studies, a strong time-on-task effect was apparent in the present data, we decided, other than preregistered, to *detrend* the present data (i.e. to remove the trend from a time series, Wang & Maxwell, 2015; Wu, Huang, Long, & Peng, 2007). A third advantage of multilevel models is that parameters can be estimated despite of missing data, for example due to unusable eye-tracking data. Finally, using a logit-link function, we can account for the dichotomous nature of our dependent variable (TRT versus TUT responses). Other than preregistered, we did not specify a model containing all eye-movement measures of interest as mind-wandering predictors due to multicollinearity problems. Instead, we specified six separate multilevel models, one for each eye-movement measure. We included crossed random intercepts[5] varying with target sentences and participants and regressed TUTs[6] (with 1 indicating TUTs and 0 indicating TRTs) on the respective eye-movement measure (Level 1), time-on-task (Level 1, variable ranging from 0 to 9, depicting the order of thought collection), and the effect-coded experimental condition (Level 2). We included time-on-task as a Level 1 predictor in order to *detrend* (see above and Wang & Maxwell, 2015) the outcome variable, to investigate the relationship between eye-movement measures and TUT occurrences above and beyond the systematic increase of TUT occurrences over time. The coefficients estimated using the *glmer* function of the R package *lme4* (Bates, Maechler, Bolker, & Walker, 2015) and the *bobyqa* optimizer for some of the models (if necessary to achieve convergence, Powell, 2009) are depicted in Table 4. Reading time, fixation count, and first-fixation duration significantly predicted TUT occurrences. Positive *b*-coefficients for these measures indicated longer reading times, longer first-fixation durations, and more fixations during TUT than during TRT episodes, as also descriptively apparent in Table 3. Within-word-regression count, between-word-regression count, and blink count did not predict TUT occurrences. In all multilevel logistic regression models, time-on-task significantly predicted TUTs (all $p$-values < .001) with the likelihood for a TUT occurrence generally increasing over time. The effect-coded experimental condition remained a non-significant predictor in all specified models (all $p$-values > .05). Thus, comparable to

[5]Adding random slopes did not improve model fits and we thus chose the more parsimonious approach.

[6] We conducted our analyses excluding all target sentences categorized as TTI episodes. In the literature, there is no common agreement on how to categorize and treat TTIs (they might reflect a thought mode between on-task and off-task thoughts). Additionally, studies examining mind-wandering using eye tracking do not differentiate between TUTs and TTIs. Using a compound TUTs + TTIs score as criterion, one more eye-movement measure proved to be a significant predictor (between-word-regression count, $b = 0.15$, Wald $Z = 2.10$, $p = .036$) compared to the analyses with only TUTs as criterion. Results for the other eye-movement measures, however, did not change.

our results on the behavioral measures, we did not find evidence that the difficulty-expectation manipulation affected eye-movement measures.

**Word-frequency effect for normal reading versus TUT episodes**

To test whether the word-frequency effect was less pronounced during TUT episodes, we compared eye-movements on high- versus low-frequency target words. We expected an interaction between word frequency (low versus high) and thought mode (TUT versus TRT), in terms of smaller differences in the eye-movement measures between high- and low-frequency target words during TUT than during TRT episodes. We employed multilevel regression models using the lme function in the R package nlme (Pinheiro, Bates, DebRoy, Sarkar, & R Development Core Team, 2010). Word frequency and thought mode (all effect-coded) entered the analyses as fixed predictors for the eye-movement measures, while allowing intercepts to vary with participant and target word identity (cf. Baayen, Davidson, & Bates, 2008). In doing so, we simultaneously controlled for participant-specific and target-word-specific variability in eye-movement measures. Again, we applied separate regression models for the prediction of gaze duration (sum of all durations of all fixations on a target word during the first pass), total viewing time (sum of all durations of all fixations on a target word including fixations following a regression), fixation count (total number of fixations on a target word), and regressions-into-target-word count (total number of regressions back to a target word). Table 5 shows estimates for main and interaction effects. For thought mode, we did not find significant main effects. A significant main effect of word frequency, with negative *b*-coefficients, represents the word-frequency effect. High-frequency target words in comparison to low-frequency target words were fixated for a shorter time (see gaze duration and total viewing time) and less often (see fixation count). Therefore, we found the typical word-frequency effect with three eye-movement measures. Regressions into target words were rare (see Figure 1) and did not show a reliable word-frequency effect.

To test whether the word-frequency effect became smaller during TUT as compared to TRT episodes, the interaction between word frequency and thought mode is crucial (see Table 5). Figure 1 additionally illustrates the interaction patterns. Descriptively, the word-frequency effect appeared to be somewhat smaller during TUT than during TRT episodes for all analyzed eye-movement measures[7]. However, only the fixation count measure yielded a significant interaction.

## Discussion

Identifying objective measures of mind wandering in addition to widely used subjective self-reports is an important endeavor, but is still a work in progress: Different authors have determined a large variety of eye-movement measures as mind-wandering indicators. However, as we reviewed in the Introduction section, there is little consensus as to the specific measures that are sensitive to attentional decoupling during mindless reading. To address this inconsistency problem, we conducted a new, high-powered eye-tracking experiment, in which we reanalyzed all previously identified mind-wandering indicators. In the present work, only three measures (reading time, fixation count, and first-fixation duration) positively predicted self-categorized mindless reading. Interestingly, aside from these single measures, the word-frequency effect was found to be less pronounced during mindless-reading than during normal-reading episodes.

In addition to our validation efforts, we tested the idea that mind wandering might be adjusted to expectancies about task demands. More precisely, we expected participants in expectation of a difficult reading comprehension test to mind wander less while reading compared to participants with low-difficulty expectations. Furthermore, we expected less mind wandering to result in better performance on the test. Indeed, we found that having fewer off-task thoughts went along with higher comprehension accuracy in the reading test,

[7] For the regressions-into-target-word count, the word-frequency effect reversed during TUT episodes. Because these regressions were very rare and the measure did not produce a reliable word-frequency effect, we refrain from strongly interpreting this finding.

but we did not find significant group differences in test performance. Additionally, although we had successfully induced opposing expectancies about task demands, the groups did not differ regarding their mind-wandering behavior. However, for participants anticipating a difficult test, higher difficulty expectations (assessed prior to the test) were associated with less mind wandering during the reading task. In this group, thoughts might have been influenced by difficulty expectations as an extrinsic motivational factor. Lacking extrinsic motivation to focus on the text, low-difficulty-expectation participants' thoughts might have been primarily influenced by intrinsic motivational factors such as interest in the topic, so that we could not find a relationship between difficulty expectancies and mind wandering. However, future research is necessary to test this assumption. Indeed, the relationship between meta-awareness and mind wandering behavior might be more complicated than assumed. In a study by Sanders, Wang, Schooler, and Smallwood (2017) directly enhancing meta-awareness of mind wandering improved task focus, but not reading comprehension, and only after a period of self-focus.

Regarding the candidate mind-wandering indicators, all three eye-movement measures that we found to be sensitive to attentional decoupling are related to reading speed. While reading mindlessly, our participants seemed to "slow down": Their fixation durations were longer and they exhibited more fixations. Consequentially, they needed more time to "read" a sentence. Changes in reading/processing time are often found to be related to mind wandering. For example, Reichle et al. (2010) also reported increased gaze durations and increased total viewing times during off-task thought episodes. Additionally, research applying the *z*-string reading paradigm (Nuthmann & Engbert, 2009; Rayner & Fischer, 1996) suggested that fixation durations are longer during mindless reading. Recently, other authors have made use of these findings and developed a machine-learned model that is able to discriminate between (self-caught) mindful and mindless reading based on eye movements (Faber, Bixler, & D'Mello, 2018). Such deceleration of cognitive processes during mind wandering episodes is not only observed for reading tasks. During word-learning tasks, off-task processing, or decoupling from the task, was found to be indexed by slower response times (Smallwood, O'Connor, Sudbery, & Obonsawin, 2007). Also, during simple go/no-go tasks, mind wandering has a prolonging effect on processing time and goes along with higher proportions of extremely long reaction times (McVay & Kane, 2012). Franklin, Smallwood, and Schooler (2011) used the relationship between mind wandering and temporal changes to develop an algorithm that successfully predicted participants' mind-wandering reports during word-by-word reading. However, the authors found participants to be speeding up during mind-wandering episodes. They discuss that this substantial difference to the results reported by Reichle et al. (2010) might be due to paradigmatic differences (word-by-word reading versus naturalistic reading). A similar speed-up is often found in the SART when participants are mind wandering (Smallwood et al., 2004). Thus, there is mixed evidence concerning the direction of the mind-wandering–processing-time relationship and further research is needed.

Franklin et al.'s (2011) algorithm is based on the idea of different sensitivities to the lexical qualities of words during mind wandering compared to on-task attention. They found people to be less affected by word features such as length and frequency during mindless reading. This idea is also supported by Reichle et al. (2010), Foulsham et al. (2013), and now, additionally, by our results. We also found a weaker word-frequency effect during mindless reading, although this reduction was only significant for the fixation-count measure. Also, from a theoretical standpoint, a reduced word-frequency effect should be a good indicator of attentional decoupling during mindless reading: While people's minds wander, they naturally disengage from the lexical input in front of them and are thus likely to ignore word features that usually affect processing time. Please note, however, that we selected target words in the present study that differed considerably in their frequency but were matched in other respects (length). Whereas this approach is optimal for demonstrating a mind-wandering-related reduction of the word frequency effect, future research is needed to determine to which degree the effects observed in this study are stimulus dependent and whether a meaningful reduction would still occur with weaker word-frequency manipulations.

In most studies examining mindless reading, probes appeared randomly during full-text reading, resulting in different to-be-analyzed text passages for each participant. Only in the Foulsham et al. (2013) study, single-sentence reading allowed for the manipulation of linguistic input. These authors compared high- and low-frequency target words during mindless versus normal reading. We applied a similar approach and also defined target sentences with specific features (high- versus low-frequency words). However, in contrast to the Foulsham et al. study, our participants were able to engage in naturalistic reading. This was achieved by letting participants' fixations on the first words of the sentence following a target sentence trigger thought probe appearances. Consequently, we could analyze predefined target sentences and words even though our participants engaged in full text reading. By keeping to-be-examined passages equal for all participants while still enabling naturalistic reading, our method combines the advantages of both full-text and single-sentence reading. We therefore recommend this method for future research.

We also recommend using multilevel models in future research. In addition to their general advantages (see Results section), they allow for the specific modeling of time-on-task effects that are often found for mind-wandering behavior (e.g., Foulsham et al., 2013; McVay & Kane, 2009; Rummel & Nied, 2017; Steindorf & Rummel, 2017). As a result, the effects of the variables of interest can be examined while controlling for systematic changes due solely to time on task. Additionally, mind-wandering research has to deal with data that are not normally distributed. This issue can be resolved by specifying responses to thought probes as a dichotomous (or categorical) dependent variable in the multilevel framework.

As we discussed in our literature review outlined in the Introduction section, the research using eye-movement measures as mind-wandering indicators is far from being coherent: Previous studies have differed regarding the lexical input that was used, the reading time, the number of participants, the analyzed areas of interest, the thought-probing procedures, and the types of statistical analyses. However, even when methods were relatively comparable, different authors analyzed different measures, found divergent effects and even effects of different directions for the same measure. We hope that the present work will aid the development of robust and stable indicators of mindless reading and we believe that it also emphasizes the importance of validation research. In general, we believe that when a field is growing as rapidly as the mind wandering area (Schooler et al., 2014; Seli, Risko, Smilek, & Schacter, 2016), researchers should not only focus on conducting more research, but also on better connecting new research with existing research. An eye-movement measure that only proves to be a valid mind-wandering predictor in one single study under unique conditions will not help the field to move forward.

Special considerations apply to the usage of within-word regressions as a mind-wandering indicator. Uzzaman and Joordens (2011) observed fewer within-word regressions during mindless reading than during normal reading, whereas we did not detect any such differences. Of course, this inconsistency might be due to differences in the specific methods applied (language, sample, font size, etc.). However, because a valid mind-wandering indicator should prove stable across such conditions (as long as the methods are fairly comparable), the present study would at least suggest that within-word regressions might be a less reliable indicator of mind wandering. Additionally, with our monitor characteristics and font size, medium sized words had a length of 2-3°. Since we applied a velocity threshold of 40 °/s to identify saccades, the detection of within-word regressions could be problematic, due to technical issues. Some words in our target sentences might have simply been too small for us to detect backwards movements, rendering the measure unreliable[8]. We would conclude that future research will be needed to determine the usefulness of within-word regressions to

[8] We thank Stephanie Huette for raising this point.

predict mind wandering, but that measures that strongly depend on technical characteristics will probably not become first choice indicators.

In sum, our study is only one other among those assessing eye-movements in the hope to detect mind-wandering indicators, wherefore our results must also be considered as preliminary. Moreover, our manipulation of TUT-levels via task-demand expectancies, which could have provided a further test of the convergent validity of the eye-movement mind-wandering indicators, was not successful. We therefore would encourage researchers who plan on conducting mind wandering studies and assessing eye movements to incorporate better convergent-validity tests. Furthermore, we would encourage them to not only report the data for indicators that turned out to be significant in their particular study, but to report results for all candidate indicators, as we did here. At the current stage of this field of research, non-significant results are just as important as significant results to determine whether eye-tracking methods are a valid tool for mind-wandering assessment.

Whereas the early studies were important for identifying eye movements as a potential mind wandering proxy, future research will also have to focus on the role methodological details (e.g., those listed in Table 1) play in the divergent results observed so far. Valid indicators of mind wandering will, in the end, have to converge across studies imposing comparable contextual demands. Additionally, plausible boundary conditions need to be identified and tested for those cases in which indicators do not converge. To achieve this goal, full transparency concerning methods and (expected) results—optimally preregistered before a study is conducted—will be key. In line with these recommendations, we hope for future research to fill the gaps that still exist concerning the validation of eye-movement measures as indicators for mindless reading.

We believe that mind-wandering research would greatly benefit if eye-movement measures were found to be robust and stable indicators of mindless reading. So far, subjective self-reports are the state-of-the-art assessment tool for mind-wandering. Not only might answers to thought probes be prone to memory driven and classification errors, but the probing procedure itself might also interrupt a participant's train of thought and impair task performance. Eye-movement measures as non-intrusive, objective mind-wandering indicators could circumvent these problems for reading tasks. Similarly, pupillometry represents a promising method for predicting TUTs during reading. Pupil dilation has been found to be higher prior to mindless reading compared to normal-reading episodes (Franklin, Broadway, Mrazek, Smallwood, & Schooler, 2013; Smallwood et al., 2011, but see Grandchamp, Braboszcz, & Delorme, 2014). Eye-movement measures and pupil dilation could bring us one step closer to real-time detection of disruptive mind-wandering behavior during reading.

Additionally, such measures may even have the potential to be used to prevent mind-wandering in situations in which it is not only disruptive, but even dangerous. Driver inattention due to mind-wandering, for example, poses a serious threat to road safety: Driving simulation studies have shown that driver behavior is indeed affected by participants' thought mode (Baldwin et al., 2017). Aside from mind wandering, fatigue is a major cause of traffic accidents and eye-state bio-feedback methods have already been tested in cars to prevent drivers from falling asleep (e.g., Devi & Bajaj, 2008). Detecting TUT episodes by means of eye-movements or pupillometry in addition to fatigue detection could be in the interest of road safety. Still, as we pointed out above, the validation of these biological markers is a work in progress. We hope that our work will drive the validation process forward so that eye movements can be used as indicators for mindless reading in the not-too-distant future.

## References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390-412. doi:10.1016/j.jml.2007.12.005

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. doi:10.1177/0956797612446024

Baldwin, C. L., Roberts, D. M., Barragan, D., Lee, J. D., Lerner, N., & Higgins, J. S. (2017). Detecting and quantifying mind wandering during simulated driving. *Frontiers in Human Neuroscience* 11, 406. doi:10.3389/fnhum.2017.00406

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed effects models using lme4. *Journal of Statistical Software, 67*, 1-48. doi:10.18637/jss.v067.i01

Cho, H.-C., & Abe, S. (2013). Is two-tailed testing for directional research hypotheses tests legitimate? *Journal of Business Research, 66*, 1261-1266. doi:10.1016/j.jbusres.2012.02.023

Christoff, K., Gordon, A. M., Smallwood, J., Smith, R., & Schooler, J. W. (2009). Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proceedings of the National Academy of Sciences, 106*, 8719-8724. doi:10.1073/pnas.0900234106

Devi, M. S., & Bajaj, P. R. (2008). *Driver fatigue detection based on eye tracking.* Paper presented at the 2008 First International Conference on Emerging Trends in Engineering and Technology, Nagpur, India.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112*, 777-813. doi:10.1037/0033-295X.112.4.777

Faber, M., Bixler, R., & D’Mello, S. K. (2018). An automated behavioral measure of mind wandering during computerized reading. *Behavior Research Methods, 50*, 134-150. doi:10.3758/s13428-017-0857-y

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist, 34*, 906-911. doi:10.1037/0003-066X.34.10.906

Foulsham, T., Farley, J., & Kingstone, A. (2013). Mind wandering in sentence reading: decoupling the link between mind and eye. *Canadian Journal of Experimental Psychology, 67*, 51-59. doi:10.1037/a0030217

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

Franklin, M. S., Broadway, J. M., Mrazek, M. D., Smallwood, J., & Schooler, J. W. (2013). Window to the wandering mind: Pupillometry of spontaneous thought while reading. *Quarterly Journal of Experimental Psychology, 66*, 2289-2294. doi:10.1080/17470218.2013.858170

Franklin, M. S., Smallwood, J., & Schooler, J. W. (2011). Catching the mind in flight: Using behavioral indices to detect mindless reading in real time. *Psychonomic Bulletin & Review, 18*, 992-997. doi:10.3758/s13423-011-0109-6

Grandchamp, R., Braboszcz, C., & Delorme, A. (2014). Oculometric variations during mind wandering. *Frontiers in Psychology, 5,* 31. doi:10.3389/fpsyg.2014.00031

Heister, J., Würzner, K. M., Bubenzer, J., Pohl, E., Hanneforth, T., Geyken, A., & Kliegl, R. (2011). dlexDB—A lexical database for the psychological and linguistic research. *Psychologische Rundschau, 62*, 10–20. doi:10.1026/0033-3042/a000029

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics, 40*, 431-439. doi:10.3758/bf03208203

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329-354. doi:10.1037/0033-295X.87.4.329

Kane, M. J., Gross, G. M., Chun, C. A., Smeekens, B. A., Meier, M. E., Silvia, P. J., & Kwapil, T. R. (2017). For whom the mind wanders, and when, varies across laboratory and daily-life settings. *Psychological Science, 28*, 1271-1289. doi:10.1177/0956797617706086

Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science, 330*, 932-932. doi:10.1126/science.1192439

Klinger, E. (1999). Thought flow: Properties and mechanisms underlying shifts in content *At play in the fields of consciousness: Essays in honor of Jerome L. Singer.* (pp. 29-50). Mahwah, NJ, US: Erlbaum.

Maas, C. J. M., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology, 1*, 86-92. doi:10.1027/1614-2241.1.3.86

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. doi: 10.1037/a0014104

McVay, J. C., & Kane, M. J. (2012). Drifting from slow to “d'oh!”: Working memory capacity and mind wandering predict extreme reaction times and executive control errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*, 525-549. doi:10.1037/a0025896

Mooneyham, B. W., & Schooler, J. W. (2013). The costs and benefits of mind-wandering: a review. *Canadian Journal of Experimental Psychology, 67*, 11-18. doi:10.1037/a0031569

Nuthmann, A., & Engbert, R. (2009). Mindless reading revisited: An analysis based on the SWIFT model of eye-movement control. *Vision Research, 49*, 322-336. doi:10.1016/j.visres.2008.10.022

Nyström, M., & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behavior Research Methods, 42*, 188-204. doi:10.3758/brm.42.1.188

Paccagnella, O. (2011). Sample size and accuracy of estimates in multilevel models. *Methodology, 7*, 111-120. doi:10.1027/1614-2241/a000029

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Development Core Team. (2010). nlme: Linear and nonlinear mixed effects models. Retrieved from https://CRAN.R-project.org/package=nlme

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology, 52*, 1-56. doi:10.1016/j.cogpsych.2005.06.001

Powell, M. J. D. (2009). The BOBYQA algorithm for bound constrained optimization without derivatives. (Cambridge NA Report 2009/NA/06, pp. 26–46). Cambridge, UK: Cambridge Centre for Mathematical Sciences.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372-422. doi:10.1037/0033-2909.124.3.372

Rayner, K., & Fischer, M. H. (1996). Mindless reading revisited: Eye movements during reading and scanning are different. *Perception & Psychophysics, 58*, 734-747. doi:10.3758/bf03213106

Reichle, E. D., Reineberg, A. E., & Schooler, J. W. (2010). Eye movements during mindless reading. *Psychological Science, 21*, 1300-1310. doi:10.1177/0956797610378686

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). `Oops!': Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia, 35*, 747-758. doi:10.1016/S0028-3932(97)00015-8

Rummel, J., & Boywitt, C. D. (2014). Controlling the stream of thought: Working memory capacity predicts adjustment of mind-wandering to situational demands. *Psychonomic Bulletin & Review, 21*, 1309-1315. doi:10.3758/s13423-013-0580-3

Rummel, J., & Nied, L. (2017). Do drives drive the train of thought?—Effects of hunger and sexual arousal on mind-wandering behavior. *Consciousness and Cognition, 55*, 179-187. doi:10.1016/j.concog.2017.08.013

Sacks, O. (2008). Fantasie nach Noten: Über musikalische Vorstellungskraft und die Macht der Töne. *Gehirn und Geist, 6*, 48 - 55.

Sanders, J. G., Wang, H.-T., Schooler, J., & Smallwood, J. (2017). Can I get me out of my head? Exploring strategies for controlling the self-referential aspects of the mind-wandering state during reading. *Quarterly Journal of Experimental Psychology, 70*, 1053-1062. doi:10.1080/17470218.2016.1216573

Schad, D. J., Nuthmann, A., & Engbert, R. (2012). Your mind wanders weakly, your mind wanders deeply: Objective measures reveal mindless reading at different levels. *Cognition, 125*, 179-194. doi:10.1016/j.cognition.2012.07.004

Schooler, J. W., Mrazek, M. D., Franklin, M. S., Baird, B., Mooneyham, B. W., Zedelius, C., & Broadway, J. M. (2014). The middle way: finding the balance between mindfulness and mind-wandering. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation* (Vol. 60, pp. 1-33). Orlando, FL: Academic Press.

Schooler, J. W., Reichle, E. D., & Halpern, D. V. (2004). Zoning out during reading: Evidence for dissociations between experience and metaconsciousness. In D. T. Levin (Ed.), *Thinking and seeing: Visual metacognition in adults and children* (pp. 204–226). Cambridge, MA: MIT Press.

Schooler, J. W., Smallwood, J., Christoff, K., Handy, T. C., Reichle, E. D., & Sayette, M. A. (2011). Meta-awareness, perceptual decoupling and the wandering mind. *Trends in Cognitive Sciences, 15*, 319-326. doi:10.1016/j.tics.2011.05.006

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

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.

Smallwood, J. (2011). Mind-wandering while reading: Attentional decoupling, mindless reading and the cascade model of inattention. *Language and Linguistics Compass, 5*, 63-77. doi:10.1111/j.1749-818X.2010.00263.x

Smallwood, J., & Andrews-Hanna, J. (2013). Not all minds that wander are lost: the importance of a balanced perspective on the mind-wandering state. *Frontiers in Psychology*, 4, 441. doi:10.3389/fpsyg.2013.00441

Smallwood, J., Baracaia, S. F., Lowe, M., & Obonsawin, M. (2003). Task unrelated thought whilst encoding information. *Consciousness and Cognition, 12*, 452-484. doi:10.1016/S1053-8100(03)00018-7

Smallwood, J., Brown, K. S., Tipper, C., Giesbrecht, B., Franklin, M. S., Mrazek, M. D., . . . Schooler, J. W. (2011). Pupillometric evidence for the decoupling of attention from perceptual input during offline thought. *PloS ONE, 6*, e18298. doi:10.1371/journal.pone.0018298

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

Smallwood, J., McSpadden, M., & Schooler, J. W. (2008). When attention matters: The curious incident of the wandering mind. *Memory & Cognition, 36*, 1144-1150. doi:10.3758/mc.36.6.1144

Smallwood, J., O'Connor, R. C., Sudbery, M. V., & Obonsawin, M. (2007). Mind-wandering and dysphoria. *Cognition and Emotion, 21*, 816-842. doi:10.1080/02699930600911531

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

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*, 389–415. doi:10.1037/aca0000046

Smilek, D., Carriere, J. S. A., & Cheyne, J. A. (2010). Out of mind, out of sight: Eye blinking as indicator and embodiment of mind wandering. *Psychological Science, 21*, 786-789. doi:10.1177/0956797610368063

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

Steindorf, L., & Rummel, J. (2017). "I should not forget the apples!"—Mind-wandering episodes used as opportunities for rehearsal in an interrupted recall paradigm. *Applied Cognitive Psychology, 31*, 424-430. doi:10.1002/acp.3328

Unsworth, N., & McMillan, B. D. (2014). Similarities and differences between mind-wandering and external distraction: A latent variable analysis of lapses of attention and their relation to cognitive abilities. *Acta Psychologica, 150*, 14-25. doi:10.1016/j.actpsy.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*, 1882-1886. doi:10.1016/j.concog.2011.09.010

Vinski, M. T., & Watter, S. (2012). Priming honesty reduces subjective bias in self-report measures of mind wandering. *Consciousness and Cognition, 21*, 451-455. doi:doi.org/10.1016/j.concog.2011.11.001

Wang, L., & Maxwell, S. E. (2015). On disaggregating between-person and within-person effects with longitudinal data using multilevel models. *Psychological Methods, 20*, 63-83. doi:10.1037/met0000030

Weinstein, Y. (2018). Mind-wandering, how do I measure thee with probes? Let me count the ways. *Behavior Research Methods, 50,* 642-661. doi:10.3758/s13428-017-0891-9

Weinstein, Y., De Lima, H. J., & van der Zee, T. (2018). Are you mind-wandering, or is your mind on task? The effect of probe framing on mind-wandering reports. *Psychonomic Bulletin & Review, 25*, 754-760. doi:10.3758/s13423-017-1322-8

Wu, Z., Huang, N. E., Long, S. R., & Peng, C.-K. (2007). On the trend, detrending, and variability of nonlinear and nonstationary time series. *Proceedings of the National Academy of Sciences, 104*, 14889-14894. doi:10.1073/pnas.0701020104

Yang, S.-n. (2006). An oculomotor-based model of eye movements in reading: The competition/interaction model. *Cognitive Systems Research, 7*, 56-69. doi:10.1016/j.cogsys.2005.07.005

Table 1. *Studies comparing eye movements during self-categorized episodes of mindless reading (MR) versus normal reading (NR).*

| | Reichle, Reineberg, & Schooler, 2010 | Smilek, Carriere, & Cheyne, 2010 | Uzzaman & Joordens, 2011 | Foulsham, Farley, & Kingston, 2013 | Frank, Nara, Zavagnin, Touron, & Kane, 2015 |
|---|---|---|---|---|---|
| *Task characteristics* | | | | | |
| *N* | 4 | 12 | 22 | 26 | 29 |
| Reading materials | *Sense and Sensibility* | Two passages of *A Short History of Nearly Everything* | 16 pages from *War and Peace* | 120 single sentences | Five chapters from *War and Peace* |
| Reading duration | 13.5 hours (average) | Up to 30 min | | 30 min | |
| Probe count | 151.5 (average) | 20 | 10 (average) | 9 (average) | 20 |
| Area of interest | 2.5 s, 5 s, 10 s, 30 s, 60 s, and 120 s intervals preceding thought probes | 5 s intervals preceding thought probes | 5 s intervals preceding thought probes | Target sentences | 3-8 s intervals preceding thought probes |
| *Eye-movement measures* | | | | | |
| Fixation count | MR = NR (first-pass fixations) | MR < NR | MR = NR | MR > NR | MR = NR (first-pass fixations) |
| Fixation duration | MR = NR (first-fixation durations) | MR = NR | MR = NR | MR > NR | |
| Between-word regression count | MR = NR | | | MR = NR | MR > NR |
| Within-word regression count | | | MR < NR | | |
| Blink count | | MR > NR | MR = NR | | MR > NR |
| Reading time | MR > NR (gaze duration & total viewing time) | | | MR > NR | MR = NR (gaze duration) |
| Word frequency effect | | | | MR < NR | |

*Note.* < and > symbols represent significant effects with $p < .05$. For example, Uzzaman and Joordens (2011) found that there were less within-word regressions during mindless compared to normal reading. For comparability reasons, only effects on eye-movement measures that appeared preceding probe-caught mind-wandering are reported from the Reichle et al. (2010) study, although the authors also examined self-caught mind-wandering. From the Frank et al. (2015) study, only effects for younger adults are reported. The authors differentiated between task-unrelated thoughts and task-related inferences, but only effects found for task-unrelated thoughts are reported here. They also advise not to use blinks as mind-wandering markers, because effects on this measure were found for only about half their participants. Uzzaman and Joordens (2011) additionally found a significant effect for a measure they called *run count*. Because it was not precisely defined, and it is not a commonly used eye-tracking measure, we did not include it as a measure of interest. MR = mindless reading; NR = normal reading.

Table 2. *Means, standard deviations, and correlations for the behavioral measures.*

| Measure | ***M*** | ***SD*** | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|---|
| 1. Comprehension accuracy (%) | ***80.87*** | ***13.18*** | -- | [-.21, .17] | [-.48, -.13] | [-.37, -.03] | [-.07, .28] | [-.34, .03] | [-.22, .17] |
| 2. Expected difficulty (1 to 10) | ***6.42*** | ***2.01*** | -.03 | -- | [-.22, .18] | [-.33, .06] | [-.06, .32] | [-.34, .01] | [-.08, .29] |
| 3. Perceived difficulty (1 to 10) | ***5.46*** | ***1.71*** | **-.31*** | -.02 | -- | [-.12, .26] | [-.24, .15] | [-.18, .23] | [-.33, .04] |
| 4. Online TUTs (%) | ***17.57*** | ***14.46*** | **-.19*** | -.15 | .08 | -- | [-.30, .11] | [.58, .79] | [-.14, .25] |
| 5. Online TTIs (%) | ***31.96*** | ***13.35*** | .12 | .15 | -.05 | -.11 | -- | [-.22, .16] | [.42, .65] |
| 6. Retrospective TUTs (%) | ***16.42*** | ***13.74*** | **-.16*** | **-.17*** | .04 | **.70*** | -.04 | -- | [-.16, .23] |
| 7. Retrospective TTIs (%) | ***27.75*** | ***13.68*** | -.03 | .11 | -.15 | .06 | **.54*** | .02 | -- |

*Note*. Values in parentheses indicate the scale of the measure. Values for the comprehension test's expected and perceived difficulty relate to a ten-point scale from *very easy* (1) to *very difficult* (10). Pearson correlations are displayed below, and bootstrapped 95% confidence intervals for the correlation coefficients are displayed above the diagonal. Numbers in the first row of the table refer to the corresponding measure in the first column. *$p \leq .050$ (one-tailed, in the predicted direction); TUT = task-unrelated thought; TTI = task-triggered intrusions.

Table 3. *Descriptive statistics for the eye-movement measures of interest for target sentences categorized as either* task-related thought (TRT) *or task-unrelated thought (TUT) episodes.*

| Eye-tracking measure | $M_{TRT}$ | $SD_{TRT}$ | $M_{TUT}$ | $SD_{TUT}$ |
|---|---|---|---|---|
| Within-word-regression count | 1.52 | 1.60 | 1.49 | 1.68 |
| Between-word-regression count | 3.17 | 2.60 | 3.16 | 2.41 |
| Blink count | 2.86 | 2.85 | 2.69 | 2.30 |
| Reading time (ms) | 9526.67 | 3624.36 | 9947.95 | 3249.09 |
| Fixation count | 33.26 | 12.16 | 34.88 | 11.26 |
| First-fixation duration (ms) | 198.46 | 32.15 | 205.68 | 31.25 |

*Note.* Values are to be interpreted sentence-wise (e.g., on average, participants exhibited 34.88 fixations per mindlessly read target sentence), except those for first-fixation duration, which are to be interpreted word-wise (e.g., on average, participants' first-fixation durations on mindfully read single words lasted 198.46 ms). In order to calculate means, we aggregated across participants and target sentences that were categorized as TRT or TUT episodes, respectively. *M* = mean*; SD* = standard deviation; TRT = task-related thought; TUT = task-unrelated thought.

Table 4. E*ye-tracking measures (Level 1) predicting TUT occurrences in separate multilevel logistic regressions also including time-on-task (Level 1) and the effect-coded experimental condition (Level 2) as further predictors. Intercepts were allowed to vary with target sentences and participants.*

| Predictor | *b* | $SE_b$ | 95% CI | Wald *Z* | *p* |
|---|---|---|---|---|---|
| Within-word regression count | 0.03 | .11 | [-0.19, 0.24] | 0.29 | .769 |
| Between-word regression count | 0.11 | .12 | [-0.12, 0.35] | 0.98 | .328 |
| Blink count | -0.11 | .13 | [-0.37, 0.14] | -0.82 | .413 |
| Reading time (ms) | 0.27 | .12 | [0.03, 0.52] | 2.24 | .025 |
| Fixation count | 0.35 | .12 | [0.12, 0.60] | 2.95 | .003 |
| First-fixation duration (ms) | 0.24 | .11 | [0.01, 0.46] | 2.11 | .035 |

*Note.* In three of the six regression models, the eye tracking measure of interest (reading time, fixation count, and first-fixation duration) positively predicted TUT occurrences. In all of the models, time-on-task significantly predicted TUTs (all $p$-values < .001) with the likelihood for a TUT occurrence generally increasing over time. The effect-coded experimental condition remained nonsignificant in all of the models (all $p$-values > .05). *SE* = standard error; CI = confidence interval.

Table 5. *Main and interaction effects of four multilevel regressions, each predicting an eye movement measure by word frequency (low versus high), thought mode (TUT versus TRT), and the interaction of both. Intercepts were allowed to vary with target words and participants.*

| Dependent variable | Predictor | *b* | $SE_b$ | 95% CI | *t* | *p* |
|---|---|---|---|---|---|---|
| | Word frequency | -45.50 | 8.38 | [-61.96, -29.04] | -5.43 | < .001 |
| Gaze duration | Thought mode | -9.49 | 8.59 | [-26.37, 7.40] | -1.10 | .270 |
| | Word frequency × thought mode | -7.94 | 8.43 | [-24.49, 8.62] | -0.94 | .347 |
| | Word frequency | -71.26 | 10.45 | [-91.79, -50.73] | -6.82 | < .001 |
| Total viewing time | Thought mode | 6.22 | 10.94 | [-15.29, 27.73] | 0.57 | .570 |
| | Word frequency × thought mode | -16.93 | 10.56 | [-37.68, 3.81] | -1.60 | .109 |
| | Word frequency | -0.24 | 0.05 | [-0.33, -0.15] | -5.23 | < .001 |
| Fixation count | Thought mode | 0.00 | 0.05 | [-0.09, 0.09] | -0.00 | .996 |
| | Word frequency × thought mode | -0.12 | 0.05 | [-0.21, -0.03] | -2.63 | .009 |
| | Word frequency | -0.01 | 0.01 | [-0.04, 0.02] | -0.53 | .597 |
| Regressions into target word | Thought mode | -0.02 | 0.01 | [-0.05, 0.01] | -1.54 | .124 |
| | Word frequency × thought mode | -0.02 | 0.01 | [-0.05, 0.01] | -1.41 | .159 |

*Note.* Three of four regression models showed a word frequency effect. That is, word-frequency predicted the eye-movement measure of interest (gaze duration, total viewing time, and fixation count). Critically, the interaction between word frequency and thought mode was a significant predictor of fixation count, indicating a reduced word-frequency effect during TUT compared to TRT episodes (see Figure 1). Regressions into target words were rare and did not show a reliable word-frequency effect. *SE* = standard error; CI = confidence interval.

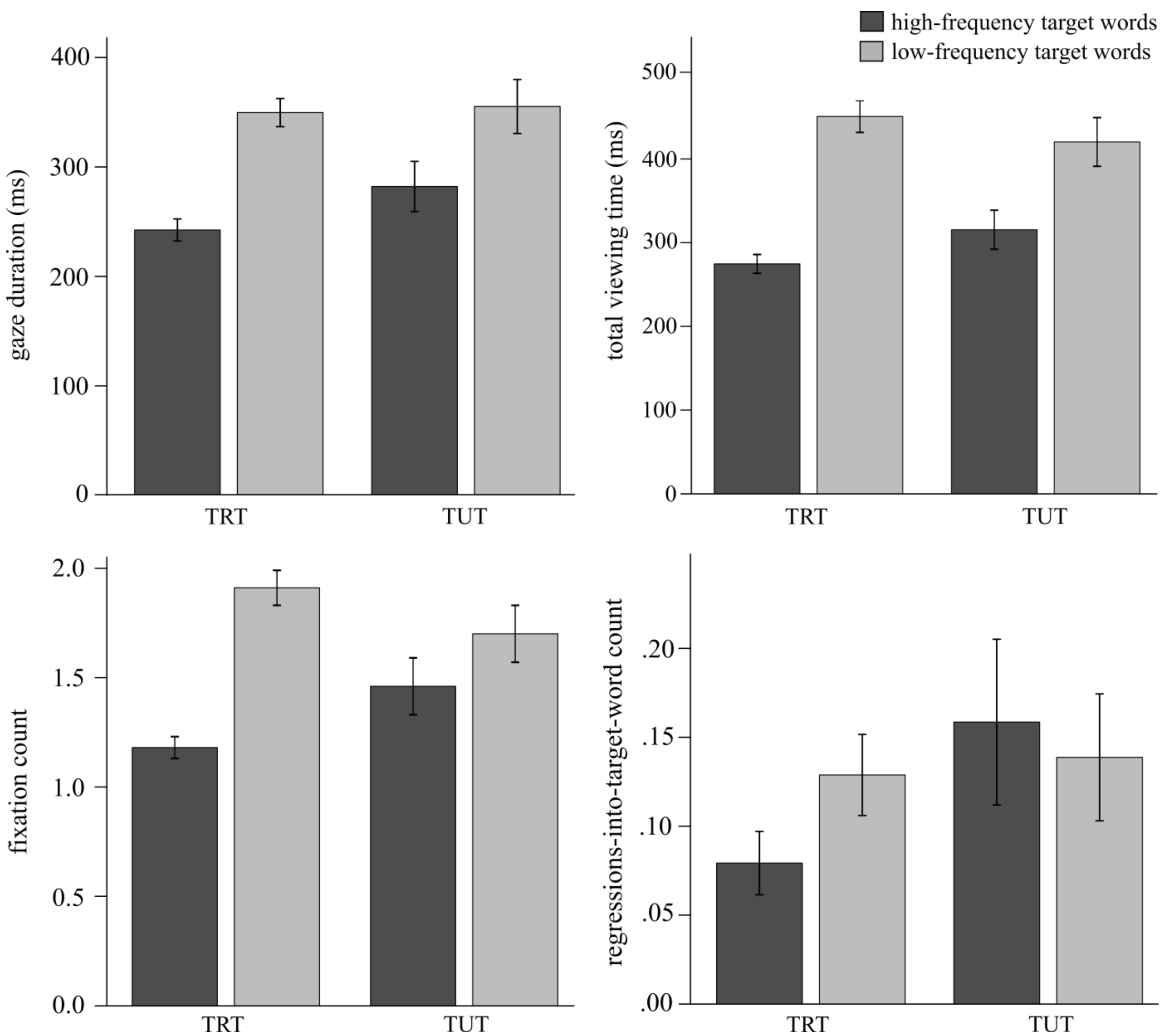


*Figure 1.*Mean values for gaze duration (duration of all fixations on a target word during the first pass), total viewing time (sum of all fixations on a target word, including fixations following a regression), fixation count (total number of fixations on a target word), and regressions-into-target-word count (total number of regressions back to a target word) for high- and low-frequency target words during TRT and TUT episodes. Higher bars for low-frequency than for high-frequency target words represent the word-frequency effect, which is descriptively smaller during TUT episodes. Because regressions into the target word were very rare and because the measure did not produce a reliable word-frequency effect, it was not interpreted further. Error bars represent standard errors of the means.

# Appendix

*Target words used to examine the word-frequency effect during episodes of mindless versus normal reading including their English translation and information on frequency and character count*

| frequency | target word (in German) | English translation | frequency | character count | mean frequency | mean character count |
|---|---|---|---|---|---|---|
| | Chanukka | Hanukkah | 4 | 8 | | |
| | Musikologe | musicologist | 1 | 10 | | |
| low | Imaginieren | imagination | 9 | 11 | 8.40 | 10.60 |
| | Pathologien | pathologies | 27 | 11 | | |
| | Frontallappen | frontal lobe | 1 | 13 | | |
| | Gehirn | brain | 2146 | 6 | | |
| | Eifersucht | jealousy | 882 | 10 | | |
| high | Strukturen | structures | 2229 | 10 | 1892.80 | 10.40 |
| | Gesundheit | health | 3773 | 10 | | |
| | Eigentümlichkeit | peculiarity | 434 | 16 | | |

*Note.* Word frequencies (type frequencies) were extracted from the *dlex* database (obtained under http://dlexdb.de/) which constitutes a corpus of the German language consisting of different text sources such as newspapers, fiction books, and scientific works (Heister et al., 2011).

**Appendix A2 – Manuscript 2**

Note:

This is a previous version of an article that has been published in its final version in *Applied Cognitive Psychology* [© John Wiley & Sons, Inc.].

Steindorf, L., & Rummel, J. (2017). "I should not forget the apples!" – Mind-wandering episodes used as opportunities for rehearsal in an interrupted recall paradigm. *Applied Cognitive Psychology, 31*(4), 424-430. https://doi.org/10.1002/acp.3328

# "I should not forget the apples!" – Mind-wandering episodes used as opportunities for rehearsal in an interrupted recall paradigm

Lena Steindorf & Jan Rummel

Heidelberg University

Authors' Note

Lena Steindorf, Jan Rummel, Department of Psychology, Heidelberg University, Heidelberg, Germany. The present research was supported by a grant from the German research foundation (DFG) to the second author (grant no: RU1996/1-1). Parts of this research were presented at the 6th International Conference on Memory, Budapest 2016.

Correspondence concerning this article should be addressed to Lena Steindorf, Department of Psychology, Heidelberg University, Hauptstrasse 47-51, 69117 Heidelberg, Germany. E-mail: lena.steindorf@psychologie.uni-heidelberg.de

**Abstract**

Mind-wandering is mostly studied for its negative effects on ongoing cognitive tasks but may be also of adaptive value. We tested the idea of mind-wandering providing opportunities for rehearsal by asking participants to study twenty grocery items for a cued recall test. After cued recall of ten items, participants were either told that the recall task was finished or that it was interrupted for another task. All participants then performed a 2-back task during which thought contents were repeatedly probed. Cued recall of the remaining items was better in the interrupted than in the finished condition and this effect was accompanied by a more efficient rehearsal strategy: Participants' thought-reports in the interrupted condition revealed a stronger and more persistent engagement in shopping-task related thoughts. Activating a relevant goal led to mind-wandering episodes being persistently used as opportunities for rehearsal revealing participants' adaptive usage of off-task thoughts.

*Keywords*: Mind-wandering, Adaptive cognition, Task-unrelated thought

# "I should not forget the apples!" – Mind-wandering episodes used as opportunities for rehearsal in an interrupted recall paradigm

As much as half of our daily thoughts can be described as being unrelated to current external events (Killingsworth & Gilbert, 2010; Klinger, 1999). You are probably not going to be able to read this article without your thoughts unintentionally trailing off every once in a while. A so called mind-wandering episode like this is most likely going to impair your reading comprehension. When thoughts are off-task, performance on a wide range of tasks like reading (Feng, D'Mello, & Graesser, 2013; Jackson & Balota, 2012; McVay & Kane, 2012), simple vigilance tasks (McVay & Kane, 2009; Smallwood et al., 2004) or even sustained attention tasks (SAT; Mrazek et al., 2012) suffers. Therefore off-task thoughts have mostly been studied for their negative effects on ongoing cognitive tasks, raising the question why people spend so much of their time mind-wandering. Given the considerable costs of off-task thoughts to concurrent tasks, it appears likely that these thoughts do have some kind of functionality. Indeed, there is recent evidence that people also sometimes intentionally engage in mind-wandering for various reasons (Seli, Risko, & Smilek, 2016). With or without intention, from an adaptive viewpoint, it seems unlikely that people would mind-wander up to 50 % of their waking hours if off-task thoughts were not functional to some extent.

Already Singer (1975) described *positive constructive daydreaming*—one out of three styles of daydreaming—as being characterized by planful, creative thought, highlighting the constructive nature of off-task thoughts. Taking a closer look at specific thought contents, it is found that off-task thoughts are often future related, concerning unfulfilled tasks or personal problems (e. g., Smallwood, Nind, & O'Connor, 2009b). Future oriented mind-wandering is not only observed under laboratory conditions but also in real life (D'Argembeau, Renaud, & Van Der Linden, 2011) and might help people to achieve personals goals or plan future actions (cf. Klinger, 1999; Mooneyham & Schooler, 2013). Besides their planful nature, off-task thoughts are also found to enhance creativity (Baird et al., 2012, but see Smeekens & Kane, 2016) and to facilitate problem solving (Ruby, Smallwood, Sackur, & Singer, 2013).

Given these potential positive effects of mind-wandering and considering the *context-regulation hypothesis* (Smallwood & Andrews-Hanna, 2013) that states that people adjust the occurrence of off-task thoughts to some extent to situational requirements, it seems advisable to engage in goal-related off-task thoughts when the situation requires and or allows it. That is, people should benefit from adjusting their thought allocation to current *and* future task goals. Especially for individuals that are aware that they will have to fulfill a certain task goal in the future, it seems advisable to maintain this task goal—even while engaging in other ongoing activities. We define goal-related mind-wandering as off-task thoughts that are used for the maintenance of future task goals. In contrast, goal-unrelated off-task thoughts concern personal issues, worries, and other thoughts unrelated to future or current task goals. Working on a current task and having in mind a future task goal, one can allocate thoughts either to the current task goal, to the future task goal (goal-related mind-wandering) or to issues unrelated to any task goal (goal-unrelated mind-wandering). In order to not sacrifice the current and future task goals, it seems advisable to prevent intrusive goal-unrelated off-task thoughts from interfering with one's goals. In line with this idea, Rummel, Smeekens, and Kane (2016) recently demonstrated that people tend to engage less in goal-unrelated off-task thoughts during an ongoing task when they have to additionally watch out for the right moment to fulfill a prospective memory intention. Also, a recent meta-analysis of neuroimaging studies (Stawarczyk & D'Argembeau, 2015) confirmed the association between mind-wandering and personal goal processing by demonstrating that both domains depend on a common set of brain regions.

There is a considerable amount of research showing that mind-wandering situationally increases due to personal states like bad mood (Smallwood, Fitzgerald, Miles, & Phillips, 2009a), nicotine cravings (Sayette et al., 2010), or acute personal worries activated via triggers (McVay & Kane, 2013). However, there are hardly any empirical demonstrations that active (versus inactive) future task goals change mind-wandering behavior (but see Rummel et al., 2016).

As a daily-life example for a situation where people will have to devote some conscious thought to a future task goal in order to maintain it despite of being occupied by other current duties, one may think of the common situation of having to do some shopping later but being engaged in another ongoing task meanwhile. In this situation, thinking about the grocery shopping list (i.e., engaging in goal-related off-task thoughts) while transcribing some text, for instance, might prove helpful later in the store. It may prevent one from forgetting to buy necessary ingredients for tonight's dinner due to mind-wandering providing opportunities for rehearsal (as proposed by Singer, 1964). In the present study, we therefore examined possible rehearsal strategies during mind-wandering episodes engaged towards the maintenance of a future shopping-task goal and their success. First, our aim was to investigate whether mind-wandering episodes are used for rehearsal and goal maintenance rather than goal-unrelated thoughts by looking at people's thought contents in the presence versus absence of a future task goal. Second, we wanted to determine the nature of rehearsal strategies by examining the persistence of rehearsal thoughts over time. Third, we assessed the success of rehearsal during mind-wandering episodes to make a point in favor of the adaptiveness of goal-related mind-wandering.

To accomplish these objectives, we asked participants to study a grocery shopping list and framed the recall phases as shopping episodes in two supermarkets. After the first recall phase (recall of half the shopping list items), we intended to manipulate mind-wandering content by activating a relevant goal in one condition: Participants in an *interrupted condition* (IC) were told that they would set out for another supermarket to buy the remaining items. While on their way, they would work on a different task and would finish the recall task afterwards. In contrast to IC participants, participants in a *finished condition* (FC) were told that they did not urgently need the remaining items, that the recall task was finished and that they would now work on a different task. We expected that with the second shopping episode still in mind, IC participants would use mind-wandering episodes more often for goal-related rehearsal than FC participants during a ten minute retention interval that was filled with a 2-back task (cf. Cohen, 2013 for a similar idea). According to our experience, the intervening 2-back task is feasible for our student population so that mind-wandering can be expected to occur during this task, on the one hand. On the other hand, we considered this task to be attention demanding enough to motivate participants to disengage from no longer relevant task goals in the FC. Also, from an applied perspective, many daily tasks impose a certain level of cognitive demands and do not allow people to unrestrictedly rehearse future task goals, like shopping lists for a later shopping trip. For these reasons, we chose the 2-back version of the *n*-back task as the intervening task accepting that it might limit the overall frequency of mind-wandering. We expected fewer goal-related thoughts (i.e., shopping-item rehearsal) for FC than IC participants during the 2-back task. Also, concerning the nature of rehearsal strategies, we expected a more persistent rehearsal pattern for IC compared to FC participants. Further, we expected these between-group differences in mind-wandering behavior to go along with better shopping item recall during the second recall phase for FC compared to IC participants.

## Method

### Participants and Design

Sixty subjects were recruited at Heidelberg University, Germany and were tested in groups. Data of 5 subjects were discarded due to poor performance in the 2-back task ($d' < 1$), resulting in $N = 55$ ($M_{age} = 22.07$, $SD_{age} = 4.01$; 84% female; 28 subjects in the FC) for all further analyses. In light of the other participants' good performance (mean $d' = 3.13$) and the moderate difficulty level of the task, we assumed that these participants did not pay sufficient attention to the 2-back task, possibly changing their mind-wandering behavior. Removal of these data did neither change the pattern nor the significance of results. We used a 2 × 2 mixed-factorial design for the investigation of off-task thoughts and 2-back performance with test-instruction condition (IC vs. FC) manipulated between participants. The time of assessment (before list encoding, after list encoding) depicted the within-participants factor. Recall performance was investigated via a 2 (IC vs. FC) × 2 (recall phase: first, second) mixed-factorial design.

**Materials**

***N*-back task and thought probes.** Participants were presented with the 2-back version of the *n*-back task. In this task, single letters were successively presented for 500 ms each with an inter-stimulus interval of 3000 ms. Participants were supposed to press a green-labeled key (*C*-key) as soon as the currently presented letter matched the one presented two trials earlier. On all other trials they were supposed to press a red-labeled key (*B*-key). Participants performed two blocks of the 2-back task. Each block contained 120 trials in total consisting of 30 non-match buffer trials (not analyzed), 50 non-match experimental trials and 40 match experimental trials. During the task, participants were interrupted by a thought probe after each sequence of twelve trials, which lasted for 42 sec., resulting in a total of ten probes per block. We made sure that a probe never appeared in the middle of a target sequence, but was always followed by 3 of the non-match buffer trials. To assess the amount of mind-wandering and the content of participants' thoughts, we used a two-step thought-probe procedure. First, participants were asked to categorize their current thoughts as being on or off task. In the second step, participants were further asked to type in a brief description of their current thought's content.

**Recall task.** For the study phase, we selected 24 grocery shopping items from the *dlex* database (Heister et al., 2011), four of which to be used as buffer items. Items were presented in random order on a white screen for 2000 ms each followed by a 500 ms inter-stimulus-interval. For the test phases we created two sets consisting of ten test items each.

Sets were matched by category (e.g. fish, fruit, vegetable), word length and word frequency. All items had different first letters which were provided as cues in the recall phases.

**Procedure**

Participants signed a consent form and answered demographic questions before receiving detailed instructions for the 2-back task and the thought probing procedure. They then performed 15 practice trials of the 2-back task which contained two thought probes. After practice, participants performed one block of the 2-back task including ten thought probes. We then asked participants to study 24 grocery shopping items for a fictitious shopping trip to the supermarket embedding a cued recall paradigm in a grocery shopping scenario. Item presentation order was determined randomly across both sets. For a first recall phase, participants were presented with the first letters from one item set (set assignment to the first and second test was counterbalanced) and asked to complete each letter in accordance with the previously studied words. After recall of the first set, critical instructions for the experimental manipulation and a 10-minute retention interval followed: All participants were told that they had left the supermarket with the recalled items, but hadn't been able to find all the items from their grocery list. IC participants were told that they would set out for another supermarket to buy the remaining items. While on their way, they would work on a different task and would finish the recall task afterwards. FC participants were told that they did not urgently need the remaining items, that the recall task was finished and that they would now work on a different task. Then, the second 2-back block (including ten thought probes) started. After this block, all participants performed a first-letter cue recall test with the study items from the second set. Finally, participants were debriefed and dismissed.

## Results

Because set order did not affect the results in any way, we collapsed across this counterbalancing factor. Performance in each 2-back block was computed in terms of *d*' scores. The amount of off-task thoughts in each 2-back block was defined as the sum of probes on which participants categorized themselves as being off task. Participants' thought contents from the open-format thought questions from the second 2-back block were rated as to whether they concerned the shopping scenario or not creating the binary dependent variable *shopping thoughts* (i.e., *goal-related off-task thoughts*) with ten measurement points. Because the shopping scenario started after the first 2-back block, shopping thoughts did not occur during the first block. Thoughts that did not concern the shopping scenario and the 2-back task were subsumed under *goal-unrelated off-task thoughts,* which concerned personal issues, worries, etc. Recall performance was defined as the number of correctly recalled items for each recall phase. Analyses of Variance (ANOVAs) were employed for the analyses of 2-back performance, self-categorized off-task thoughts, as well as recall performance. To account for the repeated thought assessment and the resulting dependencies in the data set, we used a multilevel regression model for the analysis of shopping thoughts.

**2-back task performance.** A 2 × 2 ANOVA with test-instruction condition (IC vs. FC) as between-participants and block (first, second) as within-participants factor for 2-back discrimination abilities (*d*') showed no significant difference in the 2-back task performance between blocks ($M_{(Block1)} = 3.15$, $SD_{(Block1)} = .81$; $M_{(Block2)} = 3.12$, $SD_{(Block2)} = .78$), $F(1, 53) < 1$, $p = .826$, $\eta^2_p < .01$, and test-instruction conditions ($M_{(IC)} = 3.04$, $SD_{(IC)} = .82$; $M_{(FC)} = 3.22$, $SD_{(FC)} = .60$), $F(1, 53) < 1$, $p = .356$, $\eta^2_p = .02$. The interaction was also not significant, $F(1, 53) < 1$, $p = .606$, $\eta^2_p = .01$. That is, 2-back task performance was not affected by the presence versus absence of the shopping goal[1].

[1] There was also no significant difference in the hit rates between blocks ($M_{(Block1)} = .86$, $SD_{(Block1)} = .10$; $M_{(Block2)} = .86$, $SD_{(Block2)} = .12$), $F(1, 53) < 1$, $p = .753$, $\eta^2_p < .01$, and test-instruction conditions ($M_{(IC)} = .84$, $SD_{(IC)} = .12$; $M_{(FC)} = .88$, $SD_{(FC)} = .08$), $F(1, 53) = 2.13$, $p = .150$, $\eta^2_p = .04$. The interaction remained insignificant, too, $F(1, 53) < 1$, $p = .872$, $\eta^2_p < .01$. Also, false alarm rates did not differ between blocks ($M_{(Block1)} = .04$, $SD_{(Block1)} = .04$; $M_{(Block2)} = .04$, $SD_{(Block2)} = .03$), $F(1, 53) < 1$, $p = .638$, $\eta^2_p < .01$, and test-instruction conditions ($M_{(IC)} = .04$, $SD_{(IC)} = .03$; $M_{(FC)} = .04$, $SD_{(FC)} = .03$), $F(1, 53) < 1$, $p = .547$, $\eta^2_p = .01$, and the interaction remained insignificant, $F(1, 53) < 1$, $p = .993$, $\eta^2_p < .01$.

**Mind-wandering behavior.** The amount of self-categorized off-task thoughts differed between the first ($M$ = 2.09, $SD$ = 1.90) and the second 2-back block ($M$ = 3.11, $SD$ = 2.29), $F(1, 53) = 11.76$, $p < .001$, $\eta^2_p = .18$, but not between test-instruction conditions ($M_{(IC)} = 2.70$, $SD_{(IC)} = 1.93$; $M_{(FC)} = 2.50$, $SD_{(FC)} = 1.69$), $F(1, 53) = < 1$, $p = .679$, $\eta^2_p < .01$. The interaction was not significant, $F(1, 53) < 1$, $p = .951$, $\eta^2_p < .01$. Thus, the overall amount of off-task thoughts was not influenced by the presence versus absence of the shopping goal. Overall, participants in the IC did not spend more time mind-wandering. However, they could still have used a larger proportion of their off-task thoughts for the rehearsal of shopping list items than FC participants. Figure 1 depicts the proportion of goal-related off-task thoughts versus goal-unrelated off-task thoughts during the second 2-back block. The descriptive pattern suggested that there might have been more shopping thoughts in the IC than in the FC.

To statistically secure the influence of an active shopping goal on goal-related thoughts and their persistency, we used a multilevel modeling approach. Multilevel models can account for dependencies within a data set and are thus suitable for the analysis of the dependent variable shopping thoughts which was measured repeatedly during the second 2-back block. We treated these measurements (Level 1) as being nested in participants (Level 2). To account for the binary nature of the dependent variable, we used a logit-link function to regress shopping thoughts on the effect-coded test-instruction condition (Level 2), time course (variable on Level 1 ranging from 1 to 10, depicting the time of thought assessment), and the interaction of the two. Results of this multilevel logistic regression are depicted in Table 1 and show that both test-instruction condition and time course significantly predicted the extent to which participants reported shopping thoughts. These results indicate that IC participants experienced more shopping thoughts than FC participants and that the likelihood for a shopping thought occurrence generally decreased over time. Importantly, the interaction between test-instruction condition and time course also predicted the likelihood of shopping-thought occurrence revealing differences between conditions concerning the time course of shopping thoughts, as illustrated in Figure 2. Both FC and IC participants still showed goal-related off-task thoughts at the beginning of the second 2-back block, but FC participants stopped thinking about the shopping goal after the forth thought probe, whereas participants' thought reports in the IC revealed a rather persistent mental occupation with the shopping goal.

**Recall performance.** A 2 × 2 ANOVA with test-instruction condition (IC vs. FC) as between-participants and recall phase (first, second) as within-participants factor showed that participants performed better in the first ($M$ = 6.91, $SD$ = 2.19) than in the second recall phase ($M$ = 4.25, $SD$ = 1.96), $F(1, 53) = 65.60$, $p < .001$, $\eta^2_p = .55$, and better in the IC ($M$ = 6.15, $SD$ = 1.38) than in the FC ($M$ = 5.04, $SD$ = 1.76), $F(1, 53) = 6.76$, $p = .012$, $\eta^2_p = .11$. There was a marginal interaction, $F(1, 53) = 3.05$, $p < .087$, $\eta^2_p = .05$. As evident from Figure 3, pairwise comparisons revealed no significant difference between IC ($M$ = 7.19, $SD$ = 1.88) and FC participants ($M$ = 6.64, $SD$ = 2.45) in the first recall phase, $F(1, 53) < 1$, $p = .363$, $d$ = 0.25. In the second recall phase, however, IC participants ($M$ = 5.11, $SD$ = 1.87) recalled significantly more items than FC participants ($M$ = 3.43, $SD$ = 1.69), $F(1, 53) = 12.31$, $p < .011$, $d = 0.96$.

## Discussion

We tested whether people who are in an active goal state of having to remember a shopping list would devote some of their off-task thoughts to the maintenance of the shopping items despite being occupied by another current task goal. To this end, after recall of half of the shopping list's items, participants were either told that shopping-item recall was finished or that it was interrupted for another task. Participants who considered the other task an interruption of the shopping task more intensively maintained the shopping goal over the time course of this task, as evident from the considerably higher amount and persistency of shopping-item related thoughts compared to participants for whom the shopping task was no longer active. Interestingly, participants with and without an active shopping goal reported to think about the shopping task, but participants for whom the shopping was finished stopped thinking about the future task goal early on during the following task. In contrast, participants for whom the future task goal was still active kept on thinking about the shopping list from time to time during the whole time course of the task. As expected, participants who thought they interrupted the shopping task performed better on the final recall test than participants who thought that they had already finished the shopping task. This pattern of results suggests that the increased and longer-lasting engagement in goal-related off-task thoughts may have been indeed effective in maintaining memory for the not yet recalled shopping items, thereby revealing the adaptive nature of mind-wandering about unfulfilled goals. As a daily-life example for mind-wandering benefits, it has been repeatedly suggested that thinking about a grocery shopping list while performing other tasks may prove helpful later in the store. In line with this idea, we found evidence that mind-wandering may actually prevent one from forgetting to buy necessary ingredients for tonight's dinner due to these off-task episodes providing opportunities for rehearsal.

The present experimental study was not designed to rigorously investigate correlational patterns and thus the power for correlational tests was generally very low. We aimed to manipulate mind-wandering behavior and to investigate group differences in the amount and persistency of shopping thoughts and recall performance. Although we found significant effects on all of these measures, we cannot be sure that the manipulation-induced changes in shopping thoughts caused the observed changes in recall performance. Because the distributions of the shopping-thoughts sum score deviated from a normal distribution in both groups, as indicated by Kolmogorov-Smirnov tests ($D(27) < 0.27$, $p < .001$, for the IC; $D(28) = 0.37$, $p < .001$, for the FC), the conditions for testing for a linear correlation between shopping thoughts and recall performance were not met. Therefore, and because of the low power of this test ($1 - \beta = .35$, for a medium-sized effect), we refrained from testing it here. But even if we had been able to test for this relationship it would not have been a strict test of our hypothesis that the persistency of shopping thoughts (rather than their bare amount) caused the shopping-recall advantage. In light of recent theoretical work in the intention memory literature considering attention allocation to a to-be-delayed task goal as crucial for later goal fulfillment (Smith, 2003, 2010), it seems very likely that successful shopping-goal fulfillment requires persistent mental maintenance of the goal over time. Admittedly, our design does not allow to directly test this assumption. Thus, future research needs to address this point.

The manipulation of shopping-task goals did not affect 2-back task performance implying that the changes in goal-directed off-task thoughts did not affect performance on this (moderately difficult) task at hand. This finding suggests that people are able to devote some conscious thought to active future task goals without sacrificing their concurrent task goals—probably because they engaged less in goal-unrelated off-task thoughts in the presence versus absence of the active future goal (see Figure 1). Indeed, the overall engagement in off-task thoughts was very similar in the presence and absence of the shopping-goal. Only the proportion of goal-related versus goal-unrelated off-task thoughts was influenced by the activation of a future task goal. This general idea is well in line with the context-regulation hypothesis (Smallwood & Andrews-Hanna, 2013) that suggests that the engagement and disengagement from mind-wandering depends on situational requirements (see also Rummel & Boywitt, 2014). However, the present results suggest that context adjustment is more dynamic than previously assumed. People engaged in goal-related off-task thoughts about a shopping-task goal—independent of whether they thought that the shopping task was finished or only postponed, but, whereas people without an active shopping goal stopped thinking about the shopping items relatively early on, people with a still active shopping goal maintained this goal over a longer time period despite performing another task. Also, people did not sacrifice a proportion of their current goal-related thoughts for the maintenance of the future task goal (there was no difference in the overall amount of off-task thoughts), but a proportion of their goal-unrelated off-task thoughts. It seems that with a second (future) task goal in mind, it was even more important for people to prevent intrusive goal-unrelated off-task thoughts from interfering with their task goals.

In the present study, the temporal rate of thought probe presentation was rather high (i.e., every 42 sec.) and previous research has shown that people are less likely to report mind-wandering when probes occur more frequently (Seli, Carriere, Levene, & Smilek, 2013). However, we needed to obtain a reasonable number of thought reports during a rather short period of time and our participants still reported decent levels of mind-wandering: thoughts were categorized as being off-task in about 30% of all thought probes in the second 2-back block, which is similar to the rates found in some real-life experience sampling studies (Kane et al, 2007). Also and most importantly, as both experimental groups were probed equally often, the critical group comparison could not have been influenced by the thought probe presentation frequency.

Although we found decent levels of overall mind-wandering and a significant group difference in the frequency of shopping thoughts, the latter did not occur very frequently. This might have been due to the fact that the 2-back task kept the participants readily engaged: Easier tasks without a working memory component allow for more future-oriented mind-wandering (Smallwood et al., 2009b) and one could expect to find higher levels of shopping thought reports during such tasks. Another possible explanation for the low frequency of goal-related off-task thoughts might have been the shopping task's low personal relevance. Using a predefined shopping list creates a well controllable experimental setting, but less controlled and more personally relevant materials (maybe even a self-created shopping list) might induce more goal-related mind-wandering.

The increased recall performance for the IC in the second recall phase can be considered a Zeigarnik-like effect. The *Zeigarnik Effect* (Zeigarnik, 1927) describes an advantage of remembrance of uncompleted tasks in contrast to completed tasks. Within the scope of this study's results, the observed effect could be explained by an increased mental occupation with the shopping task after the interruption. Activating the goal to later complete the task may have increased its availability and may have led to more off-task thoughts concerning the uncompleted task. This may have gone along with a more persistent occupation with the task and could thus explain the advantage of remembrance of uncompleted tasks.

To conclude, this study extends the line of research highlighting the constructive nature of off-task thoughts. Mind-wandering might help people to achieve personals goals or plan future actions (cf. Klinger, 1999; Mooneyham & Schooler, 2013), enhance creativity (Baird et al., 2012, but see Smeekens & Kane, 2016), and facilitate problem solving (Ruby et al., 2013). Additionally, we now showed that off-task thought episodes can also function in a goal-directed manner (e.g., for rehearsal of task goals that are known to become relevant in the future). A similar idea has already been pushed forward by Singer (1964) and recently by Cohen (2013) and the present results render empirical support to this idea. We conclude by pointing out that mind-wandering should not be viewed one-sidedly as only being distractive for the task at hand. We do acknowledge that performance often suffers when thoughts are off-task, but these costs of mind-wandering may be tolerable, as long as there are benefits on other—maybe more important or personally relevant—levels. If people are able to adjust their mind-wandering behavior in accordance with their current *and* future task goals, they may freely decide that remembering a shopping list in addition to one's concurrent task is either worth engaging in a particularly strong goal-related task focus (cf. Rummel et al., 2016) or more important than some of their intervening tasks and adjust their task-focus accordingly.

# 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, 23*, 1117-1122. doi:10.1177/0956797612446024

Cohen, A.-L. (2013). Attentional decoupling while pursuing intentions: A form of mind wandering? *Frontiers in Psychology, 4:693*. doi: 10.3389/Fpsyg.2013.00693

D'Argembeau, A., Renaud, O., & Van Der Linden, M. (2011). Frequency, characteristics and functions of future-oriented thoughts in daily life. *Applied Cognitive Psychology, 25*, 96-103. doi:10.1002/acp.1647

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

Heister, J., Würzner, K. M., Bubenzer, J., Pohl, E., Hanneforth, T., Geyken, A., & Kliegl, R. (2011). dlexDB - A lexical database for the psychological and linguistic research. *Psychologische Rundschau, 62*, 10-20. doi: 10.1026/0033-3042/a000029

Jackson, J. D., & Balota, D. A. (2012). Mind-wandering in younger and older adults: Converging evidence from the sustained attention to response task and reading for comprehension. *Psychology and Aging, 27*, 106-119. doi: 0.1037/a0023933

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. doi: 10.1111/j.1467-9280.2007.01948.x

Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science, 330*, 932-932. doi:10.1126/science.1192439

Klinger, E. (1999). Thought flow: Properties and mechanisms underlying shifts in content. In J. A. Singer, P. Salovey, J. A. Singer, & P. Salovey (Eds.), *At play in the fields of consciousness: Essays in honor of Jerome L. Singer.* (pp. 29-50). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

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. doi: 10.1037/a0014104

McVay, J. C., & Kane, M. J. (2012). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of Experimental Psychology: General, 141*, 302-320. doi: 10.1037/a0025250

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

Mooneyham, B. W., & Schooler, J. W. (2013). The costs and benefits of mind-wandering: a review. *Canadian Journal of Experimental Psychology, 67*, 11-18. doi:10.1037/a0031569

Mrazek, M. D., Smallwood, J., Franklin, M. S., Chin, J. M., Baird, B., & Schooler, J. W. (2012). The role of mind-wandering in measurements of general aptitude. *Journal of Experimental Psychology: General, 141*, 788-798. doi: 10.1037/a0027968

Ruby, F. J. M., Smallwood, J., Sackur, J., & Singer, T. (2013). Is self-generated thought a means of social problem solving? *Frontiers in Psychology, 4,* 962. doi: 10.3389/fpsyg.2013.00962

Rummel, J., & Boywitt, C. D. (2014). Controlling the stream of thought: Working memory capacity predicts adjustment of mind-wandering to situational demands. *Psychonomic bulletin & review*, *21*, 1309-1315. doi: 10.3758/s13423-013-0580-3

Rummel, J., Smeekens, B. A., & Kane, M. J. (2016). Dealing with prospective memory demands while performing an ongoing task: Shared processing, increased on-task focus, or both?. *Journal of Experimental Psychology: Learning, Memory, & Cognition.* Advanced online publication.

Sayette, M. A., Schooler, J. W., & Reichle, E. D. (2010). Out for a smoke: the impact of cigarette craving on zoning out during reading. *Psychological Science*, *21*, 26-30. doi: 0.1177/0956797609354059

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

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

Singer, J. L. (1964). Exploring man's imaginative world. *Teachers College Record, 66*, 165 - 179.

Singer, J. L. (1975). Navigating the stream of consciousness: Research in daydreaming and related inner experience. *American Psychologist, 30*, 727-738. doi:10.1037/h0076928

Smallwood, J., & Andrews-Hanna, J. (2013). Not all minds that wander are lost: The importance of a balanced perspective on the mind-wandering state. *Frontiers in Psychology, 4,* 441. doi:10.3389/fpsyg.2013.00441

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

Smallwood, J., Fitzgerald, A., Miles, L. K., & Phillips, L. H. (2009a). Shifting moods, wandering minds: negative moods lead the mind to wander. *Emotion*, *9*, 271. doi: 0.1037/a0014855

Smallwood, J., Nind, L., & O'Connor, R. C. (2009b). When is your head at? An exploration of the factors associated with the temporal focus of the wandering mind. *Consciousness and Cognition, 18*, 118-125. doi:10.1016/j.concog.2008.11.004

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*, 389-415. doi: 10.1037/aca0000046

Smith, R. E. (2003). The cost of remembering to remember in event-based prospective memory: Investigating the capacity demands of delayed intention performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 347-361. doi: 10.1037/0278-7393.29.3.347

Smith, R. E. (2010). What costs do reveal and moving beyond the cost debate: Reply to Einstein and McDaniel (2010). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 1089-1095. doi: 10.1037/a0019183

Stawarczyk, D., & D'Argembeau, A. (2015). Neural correlates of personal goal processing during episodic future thinking and mind-wandering: an ALE meta-analysis. *Human brain mapping*, *36*, 2928-2947. doi: 10.1002/hbm.22818

Zeigarnik, B. (1927). Über das Behalten von erledigten und underledigten Handlungen. *Psychologische Forschung, 9*, 1–85.

Table 1. *Multilevel logistic regression predicting shopping thoughts by the effect-coded test-instruction condition (Level 2), time course (Level 1), and the interaction of the two.*

| Predictor | *b* | $SE_b$ | 95% CI | Wald *Z* | *p* |
|---|---|---|---|---|---|
| test-instruction condition | 0.87 | .40 | [0.10, 1.70] | 2.16 | .030 |
| time course | -0.62 | .15 | [-0.96, -0.37] | -4.12 | < .001 |
| test-instruction condition × time course | 0.42 | .15 | [0.17, 0.76] | 2.81 | .005 |

*Note. SE* = standard error; CI = confidence interval.

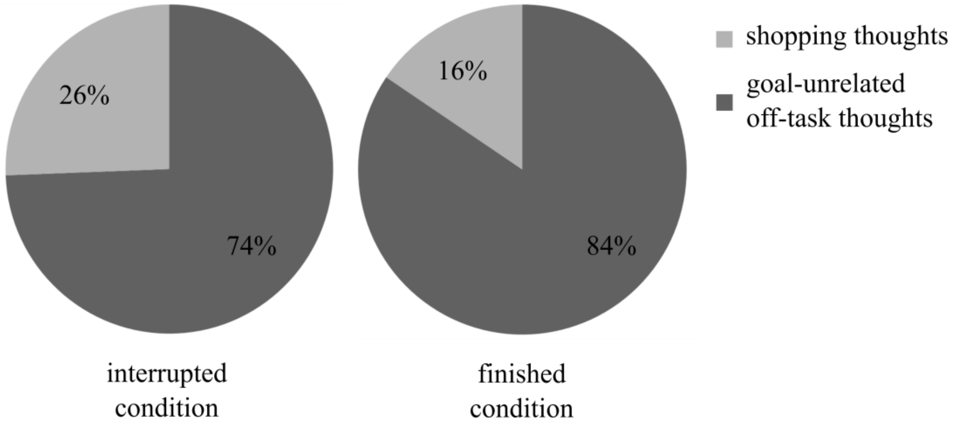


*Figure 1*. Structure of all off-task thoughts in the second 2-back block. Shown are the proportions of goal-related (shopping thoughts) versus goal-unrelated off-task thoughts for the interrupted and the finished condition.

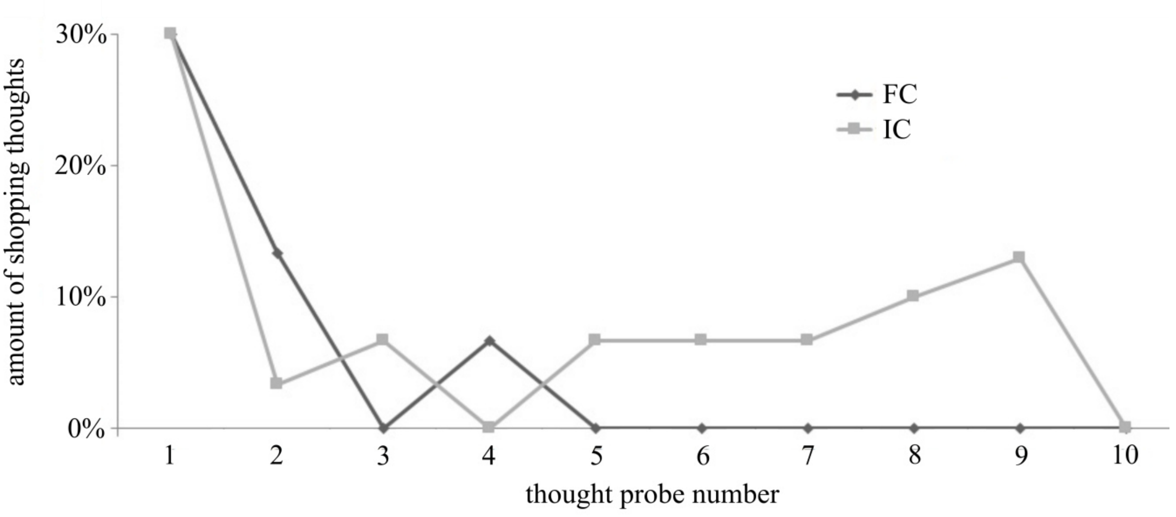


*Figure 2.* Time course of shopping-task related thoughts during the intervening 2-back task. FC = finished condition, IC = interrupted condition. Participants in both conditions showed a substantial amount of shopping thoughts at the beginning of the 2-back task that quickly decreased. However, IC participants reported some shopping thoughts almost every time they were probed, whereas FC participants stopped thinking about the shopping items after the fourth thought probe.

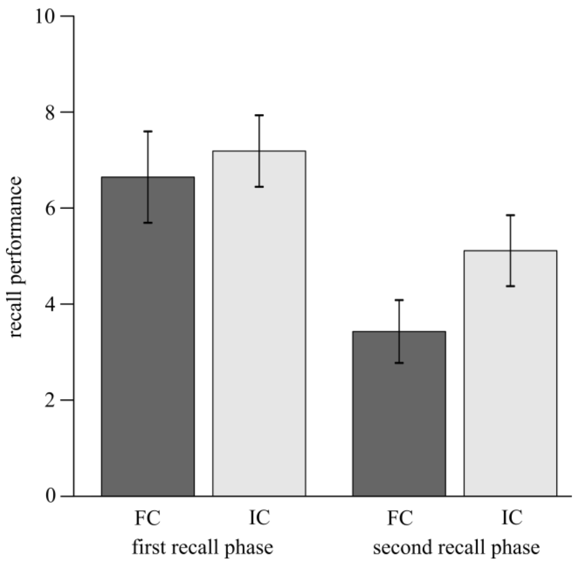


*Figure 3.* Mean recall performance values (i.e. number of correctly recalled items) for the two test-instruction conditions and recall phases. FC = finished condition, IC = interrupted condition. No differences between conditions are found for the first recall phase (before the intervening 2-back task). In the second recall phase (after the intervening 2-back task), participants in the IC recalled significantly more items than participants in the FC. Error bars depict 95% confidence intervals for the between-group comparison.

**Appendix A3 – Manuscript 3**

Note:

This is a manuscript that is currently submitted for publication. (draft version 1.8, 11th of August 2020). This version has not been peer reviewed. Please do not copy or cite without author's permission.

Steindorf, L., Boywitt, C. D., & Rummel, J. (2019). *A fresh look at the Unconscious Thought Effect: Using mind-wandering measures to reveal thought processes in decision problems with high information load.*

# A fresh look at the Unconscious Thought Effect: Using mind-wandering measures to reveal thought processes in decision problems with high information load

Lena Steindorf[1], Jan Rummel[1], & C. Dennis Boywitt[2]

[1]Heidelberg University

[2]without affiliation

Authors' Note

Lena Steindorf & Jan Rummel, Department of Psychology, Heidelberg University, Heidelberg, Germany. C. Dennis Boywitt, without affiliation. The present research was supported by a grant from the German Research Foundation (DFG) to the second and third authors (grant no: RU1996/1-1).

Correspondence concerning this article should be addressed to Lena Steindorf, Department of Psychology, Heidelberg University, Hauptstrasse 47-51, 69117 Heidelberg, Germany. E-mail: lena.steindorf@psychologie.uni-heidelberg.de

**Abstract**

Unconscious Thought Theory (Dijksterhuis, 2004) states that thinking about a complex problem unconsciously can result in better solutions than conscious elaboration. We take a fresh look at the cognitive processes underlying "unconscious" thought by analyzing data of 822 participants who worked on a complex apartment-evaluation task in three experiments. This task's information-presentation and evaluation parts were separated by different kinds of filler-interval activities, which corresponded to standard conscious-thought and unconscious-thought manipulations. Employing experience-sampling methods, we obtained thought reports during and after filler-interval engagement. Evidence concerning the existence of the Unconscious Thought Effect was mixed, with such an effect being present in the first two experiments only. In these experiments, we further found that less problem deliberation is associated with better performance on the apartment task. Interestingly, this benefit disappeared when we probed participants' thoughts during the filler interval. We suggested that explicit thought awareness diminishes the Unconscious Thought Effect.

*Keywords*: mind wandering; task-unrelated thought; Unconscious Thought Effect

## A fresh look at the Unconscious Thought Effect: Using mind-wandering measures to investigate thought processes in decision problems with high information load

### Introduction

Life is full of situations requiring decisions. Some of them are rather simple, others are rather complex. *What should I make for dinner? Which college should I go to after high school? Should I buy this washing machine or another one?* From a layman's perspective, it sounds reasonable that in such situations serious, conscious deliberation should help us make good and satisfying choices. At the same time, other people might argue that—faced with a difficult decision—one should rather sleep on it, or at least stop thinking about it for a while, to get a fresh look at the situation. Especially for complex decisions, the *Unconscious Thought Theory* (UTT, Dijksterhuis et al., 2006; Dijksterhuis & Nordgren, 2006) recommends using the latter strategy. Unconscious thought is supposed to lead to better and more satisfying decisions when choosing, for example, between four apartments, which are characterized by a multitude of attributes. In the present work, we took a closer look at thought processes during conscious- and supposedly unconscious-thought intervals by applying methods used in current mind-wandering research within a *standard* UTT paradigm. In three experiments, retrospective thought protocols as well as thought reports collected online via thought probes offered insights into the cognitive processes leading to decisions within a complex apartment-evaluation task.

Unconscious thought, which is thought or processing in the absence of conscious attention being directed towards a pending problem, was proposed as a separate form of thought distinct from and, in specific situations, superior to conscious thought (Dijksterhuis & Nordgren, 2006). In a typical UTT experiment, participants are introduced to several objects (e.g., apartments) which are characterized by a specific number of positive and negative attributes per object (e.g., "Apartment 1 has a balcony."). The objectively best object possesses a relatively high number of positive attributes, the objectively worst object a relatively high number of negative attributes. Before evaluating the objects, participants are presented with a distraction-task period or a period of conscious thought about the presented objects. Evidence for the existence of an *Unconscious Thought Effect* (UTE) comes in form of better decisions after distraction periods in complex decision-situations, that is, when objects are described by a high total number of attributes. According to the UTT, the power of the unconscious stems from its high information-processing capacity. The unconscious system is supposed to allow for large amounts of information to be integrated, whereas the conscious system suffers from a low information-processing capacity (e.g., Miller, 1956; Nørretranders, 1998). The latter refers to task-related cognitive processes that one is consciously aware of during task completion and has the advantage over the unconscious system of being rule-based and very precise (Dijksterhuis, 2004). Consequently, when faced with simple decisions, the conscious system's capacity is not exceeded and our choice benefits from rule-based cognition. We are able to consciously process all available information, which should result in the best possible decision. When faced with complex decision problems, however, its low information-processing capacity renders the conscious system less efficient because not all available information can be processed simultaneously. Here, our choice should benefit from the ability of the unconscious system to integrate a high number of decision-relevant attributes. Indeed, unconscious-thought advantages are most prevalently found for complex decision problems (e.g., Dijksterhuis et al., 2006; Strick et al., 2011).

Some assumptions of the UTT have been recently criticized and, despite many successful replications, the UTE, which states that unconscious thought improves decision making in complex problem situations, does not always replicate (e.g., Acker, 2008; Calvillo & Penaloza, 2009; Newell & Rakow, 2011; Rey et al., 2009). According to Strick et al. (2011), the debate concerning the UTE focuses on three main open questions: First, how stable and replicable is the UTE? Second, which boundary conditions are necessary for the UTE to appear? And third, what are the cognitive processes underlying periods of unconscious thought (Abadie et al., 2016, 2017; Damian & Sherman, 2013; Dijkstra et al., 2013)? Not neglecting the first two, the present work primarily addresses the third question. We were interested in participants' thoughts—unconscious as well as conscious—during distraction periods and intended to shed light on the question of which cognitive processes foster decision making or attitude formation within a standard UTT paradigm.

Schooler and Melcher (1995) suggested that incubation phases, that is phases during which a pending creative or complex problem is put aside to work on something else, cause a change of people's "mental sets": By doing so, one may get a fresh look at the situation, which eventually results in better problem solving. This view implicates a passive process of unconscious thought as something that "just happens" while a person is working on a different task. Dijksterhuis (2004) suggested that unconscious thought is rather active, as it renders mental representations more polarized as well as better organized and clustered. However, research on the UTE has been mostly *output-centered* so far. Measures such as choices, evaluations or attribute-memories have been the variables of interest used to draw conclusions concerning the cognitive processes during presumed unconscious-thought periods. For a long time, *in-the-moment* thought processes leading to specific manifestations of such output-variables have been neglected, probably because they are difficult to assess with standard cognitive methods, making it challenging to directly address the third main question (see above) raised by Strick et al. (2011). To overcome this problem, current *mind-wandering* research has been applying experience-sampling methods, which have been shown to be a valid instrument for the assessment of participants' thought contents during all sorts of tasks (see Smallwood & Schooler, 2015 for an overview). Therefore, we argue that UTT research can benefit from the employment of such methods as they have the potential to offer a "fresh look" at the cognitive processes underlying unconscious- as well as conscious-thought periods.

Mind wandering can be described as disengaged or decoupled task-attention and has been intensively studied in recent years (Mason et al., 2007; Smallwood & Schooler, 2006). Although handling many daily tasks requires our focused attention, it is a well know phenomenon that our thoughts drift off from time to time. Even though performance on the task at hand often suffers (e.g., Mrazek et al., 2012; Rummel & Boywitt, 2014), drifting thoughts also seem to be of adaptive value (Mooneyham & Schooler, 2013; Smallwood & Andrews-Hanna, 2013). For example, mind wandering towards unsolved problems or tasks may be beneficial for problem solution or task fulfillment (Baars, 2010; Steindorf & Rummel, 2017). Mind wandering is typically measured via self-reports. So-called *thought probes* interrupt ongoing tasks and ask participants to briefly describe and/or classify their current thoughts. Responses to these probes have proven to be valid (McVay & Kane, 2012) and to be good estimates of mind-wandering frequency as they do not rely on thought-awareness (Smallwood & Schooler, 2006). Finally, they are also often found to correlate with retrospective thought questionnaires (e.g., Steindorf & Rummel, 2020), which are a similar, yet distinct mind-wandering assessment method. In such questionnaires, after task completion, participants are asked to categorize the entirety of thoughts they had experienced while working on the task into several categories. In the present experiments, online as well as retrospective mind-wandering self-reports were employed to measure and quantify thought-contents occurring within a standard UTT paradigm.

Furthermore, we considered wandering thoughts as an alternative explanation for UTEs. Previous research concerning the underlying cognitive processes of complex-problem incubation suggests that either unconscious processes or short retrieval intervals during the incubation task foster post-incubation performance (Abadie et al., 2016, 2017; Damian & Sherman, 2013; Dijkstra et al., 2013; see also Sio & Ormerod, 2009). Automatically occurring, drifting thoughts concerning a still pending problem might represent such short retrieval intervals (cf. Steindorf & Rummel, 2017): Considering a typical UTT paradigm, one might argue that during the presentation of object-attribute combinations, an unsolved problem is activated within a participant's cognitive system (Watkins, 2008). This activation might lead to an increase in mental occupation with—or mind wandering towards—the problem during a period of distraction (Zeigarnik, 1927). Since mind wandering might foster problem solving (Baars, 2010), wandering apartment-thoughts during a distraction period might explain increased problem-solving performance for distracted participants. Moreover, the idea that too much deliberation can have destructive effects (Waroquier et al., 2010) as a current problem might be "thought to pieces" led us to assume that mind-wandering episodes during distraction tasks might offer just the right amount of necessary problem engagement.

Further support for our assumptions comes from the combination of findings that, during more demanding tasks, lower levels of mind wandering are reported (e.g., Rummel & Boywitt, 2014) and that more demanding distraction tasks within UTT paradigms often lead to worse problem solving compared to less demanding tasks (Abadie et al., 2013; McMahon et al., 2011; Strick et al., 2011; Waroquier et al., 2014). Focusing only on the latter finding, Strick et al. (2011) conclude that distraction tasks with high demands might compete with unconscious thoughts for resources. A different conclusion might refer to mind-wandering processes. That is, as high-demanding tasks do not leave a lot of room for mind wandering to occur, we assume that without this engagement in productive mind wandering towards the active problem, the benefit of a distraction-task period is reduced. In other words, high demands might compete with adaptive mind-wandering processes for attentional resources.

Whether high demands compete with unconscious-thought or with mind-wandering processes, both lines of argumentation suggest that during distraction tasks we allocate attentional resources towards a second ongoing cognitive process, contradicting the definition of unconscious thought as being *deliberation-without-attention* (Dijksterhuis et al., 2006). Waroquier et al. (2014) offer a solution for this dilemma by considering that attention and consciousness are not identical to each other. One can allocate cognitive resources towards the processing of specific information (attention) without being aware of the process (consciousness, or *meta-awareness* as it is often termed in the mind-wandering literature). For this reason, Strick et al. (2011) suggested replacing the term *deliberation-without-attention* with the term *deliberation-without-consciousness*. Concerning mind wandering, it is found that off-task thoughts occur with and without awareness (Smallwood & Schooler, 2006). Sometimes we know and "feel" that our minds are drifting off, sometimes we might realize that we have been pondering tonight's dinner plans only when asked about our current thoughts. However, although at times we are not aware of our thoughts at the exact moment they are occupying our minds, we are able to put these thoughts into words later, suggesting that we have nevertheless allocated attentional resources towards them. During distraction-task periods within UTT experiments, aware as well as unaware cognitive processes could foster problem solving. Our thoughts might wander towards the still active, unsolved problem, with or without awareness. Attention-demanding, but unaware wandering thoughts might be what is referred to as "unconscious" in UTT. In the present work, using self-report methods, we intended to bring wandering thoughts into awareness by directly asking participants about their current thought processes. Especially thought probes, which do not rely on thought awareness, could reveal themselves as a promising method for gaining insight into the actual attentional processes occurring during distraction-task periods. That is, participants might experience a compound of aware and unaware problem-related mind wandering, which we intend to capture and to relate to problem-solving abilities.

In the following sections, we describe three experiments, in which we hypothesize UTEs to be mirrored by changes in mind-wandering behavior. More precisely, we expected the amount of apartment-thoughts during distraction-task incubation phases (i.e., typical unconscious-thought phases) to be related to post-incubation performance. In the first experiment, we relied on retrospective mind-wandering questionnaires for a first insight into participants' thought processes during periods of distraction and conscious thought. In the second experiment, additional thought probes were employed to capture in-the-moment thoughts including unaware processes. The third experiment was conducted to take a closer look at awareness processes in UTT paradigms including thought probes. We first describe our general methods and plan of analyses before attending to the respective experiments.

## General Methods and Plan of Analyses

In all following *Methods* and *Results* sections, we report how we determined our sample sizes and all data exclusions, manipulations, and measures in the study (Simmons et al., 2012). Following the recommendations of Seli et al. (2018), we conceptualized mind wandering as task-unrelated thought and explained the concept to our participants accordingly. We named experimental conditions in which participants were instructed to think about previously presented objects during a filler interval *conscious thought conditions*. Experimental conditions in which participants worked on a distraction task during a filler interval were named *unconscious thought conditions*. These labels refer to the standard thought-mode manipulations from the UTT literature and do not imply participants' actual mode of thought, as the latter represents a to-be-examined variable in the reported experiments. Our data are available under https://osf.io/4375q/ (doi: 10.17605/OSF.IO/4375Q).

### Instruments

#### *Apartment Task*

In all three experiments, we used a German version of an apartment task[1] originally developed by Dijksterhuis (2004) to assess participants' problem solving abilities in situations with high information load. Participants of this task are presented with information about four apartments. Imagining being on apartment hunt, they are supposed to familiarize themselves

[1] We thank Arndt Bröder and his lab at the University of Mannheim, Germany for sharing their materials as well as their expertise concerning the implementation of the apartment task with us.

with and to visualize all apartments so that they will later be able to choose the best one. Each apartment is characterized by twelve attributes in total. The objectively best apartment is described by eight positive (e.g. “Apartment B has a balcony.”) and four negative (e.g. “Apartment B does not have a washing machine.”) attributes. The objectively worst apartment is described by eight negative and four positive attributes. The remaining two neutral apartments are described by six positive and six negative attributes each. For each apartment, attributes are assigned randomly from a list of twelve positive and twelve negative attributes with the only restriction being the number of positive/negative attributes. Apartment characteristics that are most essential in apartment-hunt situations (rental cost, apartment size, etc.) are not considered in the attribute list, so that they cannot overshadow other, intermediately essential, characteristics. The apartment task is typically divided into two phases, namely a *presentation* and an *evaluation* phase, which are separated by a filler interval. In the present studies’ presentation phases, the 48 apartment-attribute combinations were displayed sequentially and randomly intermixed for four seconds each, resulting in a total presentation time of 192 sec. In the later evaluation phases, participants were asked to indicate their attitude towards each of the apartments on a scale from one (*extremely negative*) to ten (*extremely positive*).

***Filler Interval Activity***

In the present experiments, two versions of the *n*-back task were used as distraction tasks within the filler interval between the apartment task’s two phases. In the *n*-back task, single letters are displayed consecutively. Participants of this task are supposed to press one key when the currently presented letter matches the one presented *n* trials earlier (target trials). For all other letters (non-target trials), they are supposed to press another key.

For the present implementations of this task, 20 different letters (*B C D F G H J K L M N P Q R S T V W Y Z*) were used and presented for 500 ms each in the center of the screen with a 300-ms inter-stimulus interval. Participants always performed one block of the *n*-back task consisting of 32 non-target and 16 target trials. The *B*-key was used as the response key for non-targets and was labeled with a red sticker. The green-labeled *C*-key was used as the target key. In the conditions with demanding distraction *n* equaled three, resulting in more letters to be constantly monitored compared to the conditions with undemanding distraction, in which *n* equaled one. Including one short introduction screen, the distraction task lasted approximately three min. Instead of working on a distraction task, the participants in the conscious thought conditions were asked to consciously think about their attitudes towards four previously presented apartments for three minutes during the filler interval.

***Retrospective Thought Assessment***

To assess the amounts of task-related and task-unrelated thoughts during the filler interval, participants of all three experiments were asked to retrospectively categorize the entirety of thoughts they had experienced during this interval into several categories using percentage scores. Participants in the conscious-thought conditions were given two response categories: (1) thoughts about the apartment task and (2) thoughts about something completely unrelated. In addition to these categories, participants in all other conditions, that is conditions including an *n*-back distraction-task, were asked to indicate the percentage of (3) thoughts about the distraction task and (4) thoughts about their performance on the distraction task. For all participants, category (2) corresponded to general off-task thoughts and was explained accordingly using multiple examples. For conscious-thought participants, category (1) corresponded to on-task thoughts. For all other participants, category (1) corresponded to a special kind of off-task thoughts and category (3) to on-task thoughts. Category (4) described task-related interferences.

**General Procedure**

The general procedure (excluding instructions and practice) is illustrated in Figure 1. At the beginning of each experiment, participants signed a consent form and provided demographic information. Afterwards, all participants received instructions for a distraction task they would have to perform later. They were then presented with the to-be evaluated apartments. The apartment task's presentation phase was followed by a filler interval that differed between conditions and experiments. Next, participants indicated their attitudes towards the previously presented apartments in the evaluation phase. Finally, participants' thought contents during the filler interval were retrospectively assessed, before participants were debriefed and dismissed. Detailed procedure descriptions for each experiment are provided below.

**General Plan of Analyses**

In all three reported experiments, we realized a conscious thought and several unconscious thought conditions, which differed with regard to the filler interval activities and thought assessment methods. We employed a consistent plan of analyses for all experiments regarding the main dependent variables, which were apartment-task performance, amounts of apartment-related thoughts during the filler interval, and amounts of task-unrelated thoughts during the filler interval. We always first ran a one-factorial ANOVA with the experimental condition as fixed factor testing for overall group differences. We then conducted follow-up analyses using Helmert contrasts, for which the first contrast always tested for an overall UTE and the following contrast(s) for the experiment-specific manipulations within the unconscious thought conditions. When necessary, we finally conducted additional simple comparisons to further disentangle significant effects.

**General Measures**

The performance on the distraction task ($n$-back) was defined in terms of the sensitivity index $d'$. The performance on the apartment task was defined as the difference in the subjective attitude values between the objectively best and worst apartment (see for example Dijksterhuis, 2004; Dijksterhuis & Nordgren, 2006). Higher values thus represent a better performance. The amounts of retrospectively reported task-unrelated and apartment-related thought during the filler interval were specified using percentage scores. Thoughts that were completely unrelated to any of the study's tasks and thus corresponded to response category (2) were considered as *task-unrelated thoughts* (TUTs) in all analyses. Thoughts that were related to the apartment task (response category (1)) were considered as *apartment thoughts* (ATs). Thoughts about the distraction task itself and the performance on this task are both distraction-task-related. Because they directly result from TUTs and ATs (($\%_{\text{distraction-task-related thoughts}} = 100\,\% - (\%_{\text{TUTs}} + \%_{\text{ATs}})$) and because the latter two were our variables of interest, we only analyzed TUTs and ATs.

## Experiment 1

We ran the first experiment to establish the apartment-task paradigm and to gain initial insights into thought processes during conscious and unconscious thought filler intervals. For this purpose, we employed a standard UTT experiment, in which one group of participants was supposed to consciously think about previously presented apartments before evaluating them. Two other groups worked on a distraction task instead during the filler interval. To additionally examine the influence of the distraction task's difficulty on thought reports as well as the apartment-task performance, we employed both an undemanding and a demanding version of the distraction task. In Experiment 1, retrospective thought reports were supposed to provide insights into participant's cognitive processes leading to the solution of the apartment problem.

### Method

#### *Participants and Design*

An a-priori power analysis with the software G*Power (Faul et al., 2007) was conducted to determine the required sample size for the planned contrast analysis central to our hypotheses that would allow to reveal medium-size effects with an $\alpha = 5\%$ and $1-\beta = 80\%$. This analysis suggested a required sample-size of at least $N = 128$. To cover for potential drop-outs, we tested a total of 153 participants at Heidelberg University, Germany in groups no larger than six. Data of five participants were excluded due to poor performance on the distraction task ($d' < 0$). Comparing these participants' performances to the non-excluded participants' good performances (mean $d' = 1.99$ for the demanding distraction task, mean $d' = 3.00$ for the undemanding distraction task), we assumed that the excluded participants did not pay sufficient attention to the distraction task, potentially changing their mind-wandering behavior and/or influencing possible unconscious thought processes. Another three participants' data were excluded due to missing values on at least one of the dependent variables of interest. Missing values resulted from participants not properly filling out the though-assessment questionnaire or the apartment evaluation. Analyses were thus executed with $N = 145$ ($M_{age} = 21.82$, $SD_{age} = 3.96$; 123 female). We employed a one-factorial design with *thought condition* being manipulated between participants: conscious thought ($n = 48$), unconscious thought with demanding distraction ($n = 49$), and unconscious thought with undemanding distraction ($n = 48$).

***Procedure***

The three experimental conditions differed regarding the filler interval activity. Accordingly, at the beginning of the experiment, the participants in the unconscious thought condition with demanding distraction received instructions for the 3-back task and were presented with practice trials. The participants in the unconscious thought condition with undemanding distraction and in the conscious thought condition (to keep the procedure equal for all conditions) read instructions for and practiced the 1-back task. Having finished the practice trials, all participants were told that they would later work on more trials of the task. For the current moment, however, they would work on a different task. All participants were instructed regarding the apartment task and the presentation of apartment-attribute combinations started. After the apartment task's presentation phase, the participants in both unconscious thought conditions worked on the respective version of the distraction task while participants in the conscious thought condition were asked to actively think about their attitude concerning all previously presented apartments. Then, in the evaluation phase of the apartment task, all participants indicated their attitude towards each of the presented apartments. Finally, they filled out the retrospective thought-assessment questionnaire.

## Results

### *Distraction-Task Performance*

The performance in the condition with a demanding distraction task (3-back, $M$ = 1.99, $SD$ = 0.70) was significantly worse than the performance in the condition with an undemanding distraction task (1-back, $M$ = 3.00, $SD$ = 0.66), $t(95) = 7.37$, $p < .001$, $d = 1.50$, reflecting the fact that the demanding task was more difficult than the undemanding one.

### *Apartment-Task Performance*

As illustrated in Figure 2a, there was a marginally significant difference between the three experimental conditions regarding the apartment-task performance, $F(2, 142) = 2.96$, $p < .055$, $\eta^2_p = .04$. Helmert contrasts indicated an UTE. That is, performance was generally worse in the conscious thought condition compared to the two unconscious thought conditions, $p < .022$. Between the two unconscious thought conditions, the apartment-task performance did not differ, $p = .451$. Thus, employing a distraction task during the filler interval, regardless of this task's difficulty, fostered participants' problem solving abilities in comparison to those participants who actively thought about the problem during the filler interval.

### *Retrospective Thought Reports*

The amount of TUTs (see Figure 2a) differed significantly between the three experimental conditions, $F(2, 142) = 50.58$, $p < .001$, $\eta^2_p = .42$. Helmert contrasts showed that the percentage of TUTs in the conscious thought condition was generally higher than in the two unconscious thought conditions, $p < .001$. Between the two unconscious thought conditions the percentage of TUTs only differed marginally, $p = .097$, with a numerically higher number in the condition with undemanding distraction (cf. Rummel & Boywitt, 2014). The amount of ATs (see Figure 2a) also varied with experimental conditions, $F(2, 142)$ =

99.59, $p < .001$, $\eta^2_p = .58$. Helmert contrasts indicated that there were more ATs in the conscious thought condition than in the two unconscious thought conditions, $p < .001$. The percentage of ATs in the undemanding unconscious thought condition was still higher than in the demanding unconscious thought condition, $p = .046$. That is, participants in the conscious thought condition thought about the apartments for roughly half of the time. During the other half, they thought about unrelated matters. Participants in both unconscious thought conditions spent the majority of their filler interval time thinking about the distraction task. However, there was still room for TUTs and ATs, especially for participants working on the undemanding task.

**Discussion**

In Experiment 1, we established the apartment task paradigm producing an UTE. This effect was independent of the distraction task's difficulty, although thought patterns differed between both unconscious thought groups. Overall, distracted participants, who reported lower levels of ATs, performed better on the apartment task. Moreover, conscious-thought participants, who showed the highest levels of apartment thought, performed worse on the apartment task. Using mind-wandering-assessment methods, we were able to demonstrate that a distraction task does not leave a lot of room for deliberation about the apartment-task problem. Hinting towards a competition for attentional resources of task-related and adaptive-mind-wandering processes, we found the lowest levels of ATs for participants in the demanding distraction condition. We had additionally expected worse problem-solving performance for participants in this condition as a result of this competition. However, participants in the demanding condition did not perform worse on the apartment task, challenging our assumptions about apartment-related mind wandering being an alternative explanation for the UTE. Still, because all participants working on any kind of a distraction task showed lower levels of ATs as well as a better apartment task performance than conscious-thought participants, the possibility remains that the unconscious thought conditions fostered “just the right amount” of ATs, which may be necessary and sufficient for a good apartment task performance.

Participants in the conscious thought condition indicated high levels of general TUTs. They only spent about half of the filler-interval time thinking about the apartments, which was their actual task. This finding might indicate that the filler interval was too long, so that participants had too much time and “thought the apartment problem to pieces.” Thus, a possible alternative explanation for our findings might be that unconscious thoughts, or fewer apartment thoughts, do not generally lead to better evaluations. Rather, too intensive conscious thought about the apartment problem may have had destructive effects on the apartment task solution.

## Experiment 2

In the second experiment, we included an additional baseline condition to better be able to interpret the effects of conscious thought manipulations. In this condition, participants did not have time to consciously (or unconsciously) think about the previously presented apartments before evaluating them. The objective was to investigate whether a high amount of ATs in the conscious thought condition would result in poorer apartment-task performance due to *overthinking* compared to no ATs in a condition without a filler interval.

Apart from including this new condition, the structure of Experiment 2 resembled the first experiment’s structure with unconscious thought conditions differing in task demands for the distraction task. Additionally, to examine thought processes in the exact moment they are happening and to capture possible unaware processes, we employed online thought probes during the filler interval. As stated in the introduction section, thought probes are frequently used in mind-wandering experiments and interrupt participants who are working on a task by asking them to briefly describe and/or classify their current thoughts’ content. Another advantage is that they do not rely on thought awareness (Smallwood & Schooler, 2006), so that they might be able to capture thought processes that participants are unaware of, which possibly are not captured as well by retrospective thought reports. Thought probes proved to be reasonably valid mind-wandering indicators (e.g., McVay & Kane, 2012) and do not interfere with performance on ongoing cognitive tasks (Wiemers & Redick, 2019). Yet, it may well be that asking participants to report on their thoughts during the filler interval might disrupt unconscious thought processes. For this purpose, we additionally employed a condition without such task-interruptions to control for possible reactive effects on apartment-task performance. Furthermore, such a condition ensures comparability with regard to Experiment 1.

## Method

### *Participants and Design*

In groups of up to six, we tested 152 participants at Heidelberg University, Germany, and 162 participants at Mannheim University, Germany, ensuring that participants had not participated in Experiment 1 and that group sizes were comparable to those in Experiment 1. Data of seven participants were discarded due to poor performance ($d' < 0$) in the distraction task (to compare, means for the non-excluded participants, $d' = 2.14$ for the demanding distraction task, $d' = 2.89$ for the undemanding distraction task). Another four participants' data were excluded due to missing values on at least one of the dependent variables of interest. Analyses were thus executed with $N = 303$ ($M_{age} = 22.78$, $SD_{age} = 4.06$; 219 female). We employed a one-factorial design with the *thought condition* being manipulated between participants (immediate evaluation ($n = 59$), conscious thought ($n = 61$), and three unconscious thought conditions: demanding distraction with thought probes ($n = 60$), undemanding distraction with thought probes ($n = 60$), undemanding distraction without thought probes ($n = 63$)).

### *Procedure*

As in Experiment 1, we employed the procedure outlined in the General Procedure section with the experimental conditions differing regarding the filler interval activity. We used a *1-back* task as the filler interval distraction task for the two unconscious thought conditions with undemanding distraction and a *3-back* task for the unconscious thought condition with demanding distraction. While working on these tasks, the participants in the conditions including thought probes were interrupted after each sequence of 12 trials (resulting in a total of four thought probes) and asked about their current thoughts. They were supposed to categorize their current thoughts' content as being *n*-back-task-related, related to their performance on the *n*-back task, related to the apartment task, or unrelated to any task in the current study. Participants in the conscious thought condition were supposed to actively think about their attitude towards all previously presented apartments during the filler interval.

Having been presented with distraction-task instructions, the apartment task's presentation phase, the filler interval activity, and the apartment task's evaluation phase in this order, participants were asked about their thoughts during the filler interval. We employed the same retrospective thought-assessment questionnaire as in Experiment 1, but for Heidelberg participants only. We decided to include this questionnaire on short notice at a time point at which data collection was already ongoing in Mannheim. The participants in the conscious thought condition filled out the version with two response options, that is, (1) thoughts about the apartment task, and (2) thoughts about something completely unrelated. The participants in all other conditions filled out the version with four response options, that is, (1), (2) as well as (3) thoughts about the distraction task, and (4) thoughts about their performance on the distraction task.

The participants in the immediate-evaluation condition worked on the individual parts of the experiment in a different order. They were asked to indicate their attitude towards each of the presented apartments directly after the presentation phase of the apartment task. That is, for these participants there was no filler interval between the apartment task's two phases. After completing the evaluation phase, they worked on the 1-back version of the *n*-back task including thought probes and filled in the retrospective thought-assessment questionnaire afterwards.

## Results

### *Immediate Evaluation Condition*

To keep the main analyses comparable between all experiments, we prepend all analyses concerning the immediate evaluation condition. Our main objective for including this condition was to test whether a large amount of ATs in the conscious thought condition would result in poorer apartment-task performance due to overthinking compared to no ATs in the immediate evaluation condition, but this was not the case. Both conditions showed a comparable apartment-task performance (see Figure 2b), $t(118) = 1.00$, $p = .321$, $d = 0.18$. Furthermore, we found that there was no significant difference between the unconscious thought conditions and the immediate evaluation condition regarding the apartment-task performance, all *p*s $\geq .149$. That is, a 3-min distraction interval within the apartment task paradigm did not result in better apartment evaluations compared to evaluations submitted right after apartment presentation.

When additionally comparing the immediate evaluation condition to the unconscious thought conditions with undemanding distraction (because participants in these conditions worked on the same task (1-back) as immediate-evaluation participants) concerning filler-interval measures, we neither found significant group differences regarding the distraction task performance (all *p*s $\geq .324$) nor did we find any significant differences regarding online as well as retrospective TUTs (all *p*s $\geq .418$). However, we found more ATs (online as well as retrospective) for participants in the unconscious thought conditions compared to those in the immediate evaluation condition, (all *p*s $\leq .043$). Participants who performed the distraction task after the apartment task was finished showed fewer ATs than participants who performed it before they made their apartment judgments. This pattern could be interpreted as a *Zeigarnik-like* effect (Zeigarnik, 1927).

***Distraction-Task Performance***

The performance on the distraction task varied between experimental conditions, $F(2, 180) = 18.15$, $p < .001$, $\eta^2_p = .17$. Helmert contrasts revealed that the performance in the unconscious thought condition with a demanding (3-back) distraction task ($M = 2.14$, $SD = 0.86$)) differed significantly from the performance in the unconscious thought conditions with an undemanding (1-back) distraction task, $p < .001$. Performance between the undemanding-distraction condition without thought probes ($M = 2.90$, $SD = 0.68$) and that with thought probes ($M = 2.82$, $SD = 0.76$) did not differ, $p = .570$. As in Experiment 1, performance on the demanding distraction task was worse than on the undemanding one. The presence of thought probes did not affect distraction task performance.

***Apartment-Task Performance***

Performance on the apartment task (see Figure 2b) varied with experimental conditions, $F(3, 240) = 3.011$, $p < .031$, $\eta^2_p = .04$. The first Helmert contrast (conscious thought versus the three unconscious thought conditions) did not indicate an UTE, $p = .515$. That is, participants who performed a distraction task during the filler interval did not generally perform better than participants who consciously thought about the apartment problem during the filler interval. Further comparing the three unconscious thought conditions with each other, we found that participants who worked on the demanding version of the distraction task achieved a similar apartment-task performance as participants who worked on the undemanding distraction task, $p = .128$. When comparing both undemanding distraction conditions, we found that participants who were not probed while working on the distraction task performed better than those who were probed, $p = .014$. Finally, employing simple contrasts to compare the conscious thought conditions with each of the unconscious-thought conditions, we found an UTE for the condition with undemanding distraction without thought probes only, $p = .026$, both other $p$s $\geq .725$ . This patter suggests that unconscious thought participants whose thoughts were not probed during the filler interval found better solutions to the apartment problem than participants who consciously thought about the apartments.

***Retrospective Thought Reports***

As in Experiment 1, we analyzed TUTs and ATs as reported in the retrospective questionnaires (see Figure 2b). Because we had retrospective thought data available for the Heidelberg participants only, we ran the analyses with $N = 116$. The amount of TUTs differed significantly between the experimental conditions, $F(3, 112) = 40.47$, $p < .001$, $\eta^2_p = .52$. Helmert contrasts revealed that conscious thought participants reported more TUTs than participants in the unconscious thought conditions, $p < .001$. Participants who had worked on the demanding distraction task showed fewer TUTs than participants who had worked on the undemanding distraction task, $p = .006$. Undemanding distraction-task participants who had received thought probes showed similar levels of TUTs as participants who did not receive any thought probes, $p = .407$.

Regarding the amount of ATs, there was a significant difference between the experimental conditions, $F(3, 112) = 25.92$, $p < .001$, $\eta^2_p = .41$. Helmert contrasts showed that more ATs were reported in the conscious thought condition than in the unconscious thought conditions, $p < .001$. Participants who had performed the demanding distraction task showed fewer ATs than participants who had performed the undemanding distraction task, $p = .014$. Moreover, undemanding-distraction participants, who had received thought probes while working on the distraction task, showed similar levels of ATs as participants who did not receive thought probes, $p = .732$.

These findings suggest that conscious thought participants had followed our instructions to consciously think about the apartments. However, they only did so for roughly 40 percent of the time. The high demands of the 3-back task almost entirely kept participants in the demanding unconscious thought condition from thinking about the apartments and unrelated matters whereas participants in both undemanding unconscious thought conditions still thought about the apartments and other issues from time to time. Thought probes had no influence on reported thoughts.

***Online Thought Reports***

The amounts of online reported TUTs and ATs were defined as the sum of thought probes in which participants self-categorized their thoughts as being task-unrelated or apartment-related, respectively. Unconscious-thought participants who had performed the demanding 3-back version of the distraction task ($M$ = 0.42, $SD$ = 0.59) reported significantly fewer TUTs than unconscious-thought participants who had performed the undemanding 1-back version ($M$ = 1.17, $SD$ = 1.04), $t(118) = -4.84$, $p < .001$, $d = -0.88$. The same pattern was found for ATs, with a fewer ATS for participants who had worked on the demanding version ($M$ = 0.18, $SD$ = 0.43) compared to those who had worked on the undemanding version of the distraction task ($M$ = 0.46, $SD$ = 0.67), $t(118) = -2.90$, $p = .004$, $d = -0.53$. Overall, patterns of online thought reports mirrored retrospective thought reports patterns. Even numerically, percentages of TUTs and ATs were very similar whether measured by online or by retrospective methods[2].

**Discussion**

In Experiment 2, we found an UTE only for participants who did not receive thought probes while performing the distraction task during the filler interval. These participants performed better in the apartment-task than participants who consciously thought about the apartments during the filler interval. They also retrospectively reported less apartment-related thoughts compared to conscious thought participants, replicating the findings from Experiment 1. The second experiment's results also ruled out the possibility that conscious

---

[2] Calculating percentages from the total values yielded 10.50 % TUTs and 4.50 % ATs out of 100% for participants in the condition with demanding distraction including thought probes. In the condition with undemanding distraction, 29.25 % TUTs and 11.50 % ATs were reported.

thought participants might overthink the apartment problem, as they did not perform worse than participants who were not given time to consciously or unconsciously think about the apartments before evaluating them.

In Experiment 2, we validated the retrospective thought reports employed in Experiment 1. Thought probes provided similar thought descriptions as retrospective questionnaires. This finding suggests that after and during task completion, participants seem to be well aware of their recent thought processes. Because thought probes were found to be valid mind-wandering-frequency indicators (McVay & Kane, 2012) which do not rely on thought awareness because they capture participants' thoughts in the moment they are happening (Smallwood & Schooler, 2006), we conclude that we might have captured a compound of aware as well as unaware wandering thoughts. A distinction between the two kinds of mind-wandering assessment might be that retrospective reports were not intrusive, due to them capturing mind wandering after distraction-task completion, that is, without interfering with any thought processes during the task. Thought probes, however, might have altered participants' thought experiences during the distraction task by bringing the unaware portions of the thought compound into awareness.

Unexpectedly, we found that the UTE disappeared when we employed thought probes during the filler interval, leading us to assume a detrimental nature of such probes to the processes producing the UTE. It has already been found that changing thought-probe characteristics within mind-wandering experiments can lead to differences in results (Robison et al., 2019; Seli et al., 2013; Weinstein et al., 2018), which made it reasonable to assume that the mere presence of thought probes might have interfered with (unconscious) thoughts for at least two reasons (Steindorf et al., under review): One reason could be that thought probes may have interrupted the ongoing task and thereby ongoing (unconscious) thought processes. Alternatively or additionally, thought probes may have made participants more aware of their current states of thought during the distraction task, whereas an absence of thought awareness might be a necessary criterion for an UTE to appear. To test these two competing assumptions and to further replicate the negative association between AT-levels and apartment-task performance, we conducted a third experiment.

## Experiment 3

In the third experiment our aim was to, once more, replicate the UTE and its negative association with the number of ATs. For this reason, we included the same conscious thought condition as well as the same unconscious thought condition (undemanding distraction task without thought probes) as in the first and second experiments. To take a closer look at the effect of thought probes on thought processes as well as on apartment-task performance, we included one condition with an undemanding distraction task and thought probes (see Experiment 2) and one new condition in which participants were only interrupted during the undemanding distraction task, but not asked about their current thoughts. If mere interruption was responsible for the lack of an UTE when including thought probes as found in Experiment 2, the UTE should be absent in this condition. If, however, thought awareness was a necessary condition for the effect to vanish, we would expect a better apartment-task performance in this condition compared to both the condition including regular thought probes and the conscious thought condition. A lack of thought awareness as a necessary condition for the UTE would further support the assumption of a deliberation-without-consciousness effect (Strick et al., 2011). Thought probes might add consciousness/awareness to the deliberation part, diminishing its beneficial effect on problem solving performance.

### Method

#### *Participants and Design*

In groups of up to six, we tested 289 participants at Heidelberg University, Germany, and 108 participants at the University of Mannheim, Germany. To be able to calculate correlations between AT-levels and apartment-task performance within experimental conditions, we substantially increased the group sizes for this experiment. Participants of

Experiment 3 had not taken part in Experiment 1 and 2. Data of one participant were discarded due to poor performance ($d' < 0$) in the distraction task (to compare, mean for the non-excluded participants, $d' = 3.07$). Another five participants' data were excluded because they were handed out no or a wrong retrospective thought-assessment questionnaire. Yet another 17 participants' data were excluded due to missing values on at least one of the dependent variables of interest. Analyses were thus executed with $N = 374$ ($M_{\text{age}} = 21.64$, $SD_{\text{age}} = 3.19$; 292 female). We employed a one-factorial design with the *thought condition*[3] being manipulated between participants. We had a conscious thought ($n = 96$) and three different unconscious thought conditions with undemanding distraction: one without thought probes ($n = 92$), one with thought probes ($N = 94$), and one with trivia probes ($n = 92$).

***Procedure***

We adhered to the general procedure and implemented the differences between experimental conditions within the filler interval as follows: We employed the *1-back* version of the *n*-back task as the filler interval distraction task for all three unconscious thought conditions. The two conditions with and without thought probes were equivalent to those employed in Experiment 2. Instead of being presented with thought probes, participants in the trivia-probe condition were presented with a trivia question after each sequence of 12 trials

[3] While running the experiment, we added a fifth condition and collected data from 42 additional participants in this condition with *modified thought probes*. Only slightly deviating from the procedure with regular thought probes, participants in the condition with modified thought probes were supposed to categorize their current thoughts' content as being task-related (related to the 1-back task), related to their task performance, or task-unrelated. That is, we did not mention the apartment task in any of the response options to the modified thought probes. This group's data (apartment-task performance ($M = 3.00$, $SD = 0.61$), TUTs ($M = 18.64$, $SD = 16.54$), ATs ($M = 9.83$, $SD = 11.84$)) will not be further considered, because it did not deviate from the regular online thought-probe condition and we thus would not have gained any new insights from adding it to our design.

(as was the case for the thought probes). Each trivia question had four response options, only one of which was correct (e.g., *Who wrote the fantasy novels "Lord of the Rings"? (1) John Ronald Reul Tolkien (2) Joanna Kathleen Rowling (3) Pete Johnson (4) Jeff Kinney*). As in the previous experiments, participants in the conscious thought condition were supposed to actively think about their attitude toward all previously presented apartments during the filler interval.

Having been presented with instructions, the apartment task's presentation phase, the filler interval activity, and the apartment task's evaluation phase in this order, participants were asked about their thoughts during the filler interval. We employed the same retrospective thought-assessment questionnaire as in Experiments 1 and 2. The participants in the conscious thought condition filled in the version with two response options, that is, (1) thoughts about the apartment task, (2) thoughts about something completely unrelated. The participants in all other conditions filled in the version with four response options, that is, (1), (2), (3) thoughts about the distraction task, and (4) thoughts about their performance on the distraction task.

## Results

### *Distraction-Task Performance*

All unconscious thought conditions featured the same 1-back distraction task in Experiment 3. As expected, there was no significant difference between the unconscious thought conditions concerning the performance on the distraction task, $F(2, 275) = 1.26$, $p < .29$, $\eta^2_p = .009$.

### *Apartment-Task Performance*

There was no significant main effect of the experimental condition for the performance on the apartment task (see Figure 2c), $F(3, 370) = 0.38$, $p = .765$, $\eta^2_p = .00$. Simple contrasts further indicated that performance in each unconscious thought condition was comparable to the performance in the conscious thought condition, all *p*s $\geq .343$.

***Retrospective thought reports***

As in Experiments 1 and 2, we analyzed TUTs and ATs as reported on the retrospective questionnaires (see Figure 2c). The amount of TUTs during the filler interval varied between experimental conditions, $F(3, 370) = 41.94$, $p < .001$, $\eta^2_p = .25$. The first Helmert contrast indicated that the percentage of TUTs was higher in the conscious thought condition compared to all three unconscious thought conditions, $p < .001$. Further comparing the unconscious thought conditions with each other, we found that participants who had received thought probes reported higher levels of TUTs than participants who had received trivia probes or no thought probes, $p < .001$. The latter two conditions did not differ from each other concerning the amount of TUTs, $p = .649$.

The amount of ATs also varied between experimental conditions, $F(3, 370) = 228.70$, $p < .001$, $\eta^2_p = .65$. Helmert contrasts indicated higher AT levels in the conscious thought compared to all other conditions, $p < .001$, but no differences between thought-probed, trivia-probed and unprobed unconscious-thought conditions, all *ps* $\geq .122$.

***Online Thought Reports***

Thought probes were employed in one condition only. Participants reported numerically similar levels of TUTs ($M = 1.22$, $SD = 1.01$) and ATs ($M = 0.38$, $SD = 0.55$) as in the corresponding condition in Experiment 2.

***Correlational Analyses***

To more directly test for a relation between ATs during the filler interval and later apartment-task performance, we correlated these measures within conditions. Retrospectively reported ATs did not correlate with the apartment-task performance in any of the experimental conditions, all *p*s $\geq .242$.

**Discussion**

In Experiment 3, we did not replicate the UTE, which we had previously found in both Experiments 1 and 2. In addition, comparisons of apartment-task performance between the unconscious thought conditions were not conclusive. We did not replicate the detrimental effect that thought probes had had in Experiment 2. Consequently, we cannot draw any conclusions concerning a lack of thought awareness as a necessary criterion for the UTE to appear. However, thought probes had an influence on retrospectively reported general TUTs, with more wandering thoughts found in participants who had been explicitly asked about their state of thought compared to participants whose thoughts had not been probed. This influence of thought probes goes beyond mere interruption, because participants who were merely interrupted did not show such increased levels of TUTs. One could speculate that the online thought probes made participants more aware of mind wandering instances during the distraction task which, in turn, resulted in higher levels of retrospectively reported TUTs.

**Joined Analysis of the Unconscious Thought Effect**

Having employed the same conscious thought and the same unconscious thought (undemanding distraction without thought probes) condition in all three experiments allowed us to collapse the data for these conditions to conduct a joined analysis of the UTE with $N$ = 408 ($n_{\text{conscious}}$ = 205, $n_{\text{unconscious}}$ = 203). We ran a 2 (experimental conditions) x 3 (experiments) ANOVA for the apartment-task performance, which revealed no significant main effect of experiment, $F(2, 402) = 1.23$, $p = .295$, $\eta^2_p = .01$, but a significant effect of experimental condition, $F(1, 402) = 4.33$, $p = .038$, $\eta^2_p = .01$, suggesting that, overall, there was an UTE: Participants in the unconscious thought condition ($M$ = 2.73, $SD$ = 2.69) performed better on the apartment task than participants in the conscious thought condition ($M$ = 2.33, $SD$ = 2.66), which equates to a small effect with a Cohen's $d$ of .15. Notably, the interaction was significant, $F(2, 402) = 3.42$, $p = .034$, $\eta^2_p = .02$, bolstering the cross experimental observation that the UTE was, at least tentatively, present in Experiments 1 and 2, $p$s ≤ .09, but not in Experiment 3, $p = .345$.

## General Discussion

### Unconscious Thought Effect

Making tough decisions without any cognitive effort sounds like a good deal. Because of this obvious appeal, the UTT has been extensively studied. A meta-analysis by Strick et al. (2011) identified the UTE as a real effect, which is, however, moderated by many factors such as distraction-task features, problem-presentation features, and filler-interval length. A later meta-analysis by Nieuwenstein et al. (2015) heavily criticized the UTT, stating that there exists no support for its notions. Adding to this ongoing discussion, we ran three experiments with the goal to get a fresh look at the UTE and its underlying cognitive processes. Overall, in our joined analysis, we found an UTE of a small effect size. However, even though we employed similar methods in all three experiments, our findings were still mixed. An UTE was present in Experiments 1 and 2, but absent in Experiment 3. The possibility remains that due to Type II errors this is a natural occurrence when running nearly the same experiment repeatedly.

However, furthermore speaking against the UTT, we found that unconscious thought participants did not produce better apartment evaluations compared to immediate-evaluation participants in Experiment 2. As for example Waroquier et al. (2010) stated, it might be just as good to trust your first intuition as to think unconsciously. Also, thinking consciously did not benefit (nor harm) apartment evaluations when taking immediate evaluations as a baseline, suggesting that it might indeed be the most efficient strategy to „just trust your immediate gut feeling“.

### Thought Processes Within a UTT Paradigm

Current mind-wandering research has been applying self-report methods such as online thought probes or retrospective questionnaires. These instruments allow scientists to assess participants’ internal thought processes, which is the reason why we implemented them within a standard UTT paradigm. We believed that an insight into participants’ thoughts during unconscious- as well as conscious-thought periods might help us to better understand the processes leading to decisions within a complex-problem scenario.

Across all three experiments, we found that conscious-thought instructions indeed led to considerably higher levels of mental occupation with a previously presented evaluation problem. Still, participants did not use all of the available time for conscious problem deliberation. Indeed, they spent roughly about half of the time thinking about the problem and the other half thinking about unrelated matters. It appeared as if the deliberation time had been too long so that people's minds started wandering. However, Experiment 2 ruled out the possibility that participants in a conscious condition would overthink the problem at hand, possibly deteriorating evaluation quality: Participants in the conscious-thought condition performed similarly well as participant in the immediate-evaluation condition. However, this also implies that high amounts of conscious thought about the evaluation problem did not lead to better evaluations compared to those formed only by first impressions, so that one could go so far as to describe conscious thought as unnecessary. In addition, the results of Experiments 1 and 2 revealed that conscious-thought participants showed worse performances evaluating the apartments than unconscious-thought participants (unless thought probed), who showed significantly lower levels of problem-related thought. Only taking into account these two experiments, one could indeed argue that lower levels of problem deliberation foster higher quality evaluations, as is assumed by the UTT. However, the results from Experiment 3 put this notion into perspective, as evaluation performance did not differ between groups, although the extent of problem deliberation differed.

In a survey concerning real life purchase decisions, Dijksterhuis et al. (2006) asked participants how much they had thought about a product they had recently bought. For complex products, the authors found that a higher amount of conscious product-thought was associated with lower satisfaction with the product. To directly test for such an association within our paradigm, we correlated the amount of problem-related thought with problem-solution quality within the experimental conditions in Experiment 3. These correlations were around zero (and not significant). That is, we did not find evidence for a relationship between the amount of ATs and decision performance within conditions on an individual level. It may still be, however, that the individual variations in ATs within each condition were just too small to affect decision performance.

In the present work, we constantly found that distraction-tasks with high demands did not leave as much room for mind wandering, problem-related or unrelated, as tasks with low demands. Such effects of task demands are found to be stable within the mind-wandering literature (e.g., Rummel & Boywitt, 2014) and suggest that task demands compete with wandering thoughts for attentional resources. Concerning a UTT paradigm, our results suggest that demanding filler-interval tasks occupy attentional resources to a higher degree than undemanding tasks, leading to fewer conscious problem-related thoughts. Before, we considered wandering thoughts as an alternative explanation for UTEs and argued that high filler-task demands would compete with adaptive mind-wandering processes for attentional resources. Further, we argued that without engagement in productive mind wandering towards the pending evaluation problem, the benefit of a distraction-task period would be reduced. Although the results within the different unconscious-thought conditions mirrored the first part of this line of reasoning, there was no evidence for a connection between higher amounts of problem-related mind wandering and better evaluations, qualifying our alternative explanation for UTEs. Yet, because working on any kind of distraction resulted in lower amounts of problem-related thoughts as well as better evaluations than engaging in conscious problem thought (at least in Experiments 1 and 2) in general, the possibility remains that “less is better,” or that there is “just the right amount” of problem-related thought which is necessary and sufficient for good decision making. Research showing that self-paced conscious thought periods are shorter than experimenter-paced conscious thought periods whilst also leading to better decisions supports this assumption (Payne et al., 2008).

Furthermore, the question remains whether what we called ATs in all our experimental conditions is qualitatively the same across conditions. Instructing participants to consciously think about a solution to the apartment problem might lead to other kinds of thought than asking them to work on a letter-task. Conscious-thought participants might have actively tried to engage in various strategies such as remembering attributes and weighting them, which might not be the best strategy within a complex decision situation given the conscious' system's low information-processing capacity. By contrast, while working on a distraction-task, ATs might have been of a completely different nature, possibly more focused on a holistic visualization of or a feeling evoked by a respective apartment. Such potential differences of qualitative nature should be addressed in further research, which could possibly employ qualitative methods and more detailed thought reports.

**Effects of Thought Awareness**

Another issue which we addressed in the present work was whether UTEs are really the result of deliberation without attention (Dijksterhuis et al., 2006). Strick et al. (2011) already suggested replacing this term with the term deliberation-without-consciousness, focusing on the distinction between attention and consciousness, or rather awareness. Participants of an unconscious-thought experiment might allocate attentional resources towards the active problem during a filler interval, without being aware of this process. The same is true for mind wandering in general, which is found to occur with and without awareness (Smallwood & Schooler, 2006). By applying thought probes in Experiment 2, we aimed to capture both aware as well as unaware thoughts in order to relate them to problem-evaluation quality. The results indicated that self-reports from online thought probes mirrored those from retrospective questionnaires, thereby validating each other. After and during task completion, participants seem to be well aware of their recent thought processes, when asked. Although both mind-wandering assessment methods produced similar estimates on thought variables, evaluation performance varied between the two conditions which differed in nothing but the mind-wandering assessment method. A distinction between the two kinds of assessment might be that thought probes are intrusive and might bring thought processes into awareness during the task. A lack of thought awareness might, however, be a necessary criterion for an UTE to appear, which would explain the increase in evaluation performance for the condition without thought probes only. However, results from Experiment 3, which was supposed to test this assumption, were inconclusive.

Depending on the choice of mind-wandering assessment method, results of studies applying thought probes have been found to fluctuate (Seli et al., 2013; Weinstein, 2018; Weinstein et al., 2018). Given that changing probe characteristics can lead to considerable differences in results, one might argue that the mere presence of probes produces similar or even larger discrepancies. As we recently found in our lab (Steindorf et al., under review), probing participants' thoughts during an incubation interval within a creativity task resulted in fewer creativity-task-related thoughts (reported retrospectively) compared to not applying any probes or applying trivia probes (cf. present Experiment 3). Because trivia probes interrupt participants just as thought probes do, this effect cannot be attributed to mere task-interruption. We interpreted the findings as an awareness-effect: Thought probes might have made participants more aware and more cautious of their mind-wandering behavior, thereby changing the experience itself. Also, in the present work's third experiment, thought reports were affected by differences in mind-wandering assessment methods. Participants who were thought-probed retrospectively reported more general TUTs compared to those who were trivia-probed or not probed at all. Asking people about their current thoughts might either change their in-the-moment mind wandering behavior or their recollection of mind-wandering instances when asked after task completion. Further research will be needed to investigate the potentially intrusive nature of thought probes for complex decisions in more detail.

To gain further insights into thought-awareness processes during UTT experiments, further research could rely on self-caught mind-wandering assessment. Employing such methods, a researcher asks participants to monitor their awareness of mind-wandering instances (Cunningham et al., 2000). Unlike probe-caught mind wandering, it requires participants to be aware of their thought experiences (Smallwood & Schooler, 2006). Employing both measures within one and the same experiment could help disentangle actual mind-wandering instances and awareness of such instances (e.g., Sayette et al., 2009).

**Conclusion**

Across three experiments, we examined the nature of thoughts during UTT experiments. We relied on retrospective mind-wandering questionnaires as well as on online thought probes to gain insights into participants' cognitive processes during distraction and conscious-thought periods. Although we found evidence for the existence of an UTE in the first two experiments, results of the third experiments were inconclusive. In the first two experiments, participants who reported fewer problem-related thoughts during a filler interval showed better problem-solving abilities later, however only when their thoughts were not probed online during the distraction task. We took a closer look at thought-awareness processes in Experiment 3 and found that thought reports were influenced by the applied mind-wandering assessment method. Problem solving, however, was unaffected.

## References

Abadie, M., Waroquier, L., & Terrier, P. (2013). Gist memory in the unconscious-thought effect. *Psychological Science, 24*, 1253-1259. https://doi.org/10.1177/0956797612470958

Abadie, M., Waroquier, L., & Terrier, P. (2016). Information presentation format moderates the unconscious-thought effect: The role of recollection. *Memory, 24*, 1123-1133. https://doi.org/10.1080/09658211.2015.1070179

Abadie, M., Waroquier, L., & Terrier, P. (2017). The role of gist and verbatim memory in complex decision making: Explaining the unconscious-thought effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*, 694-705. https://doi.org/10.1037/xlm0000336

Acker, F. (2008). New findings on unconscious versus conscious thought in decision making-additional empirical data and meta-analysis. *Judgment and Decision Making, 3*, 292 - 303.

Baars, B. J. (2010). Spontaneous repetitive thoughts can be adaptive: Postscript on "mind wandering". *Psychological Bulletin, 136*, 208-210. https://doi.org/10.1037/a0018726

Calvillo, D. P., & Penaloza, A. (2009). Are complex decisions better left to the unconscious? Further failed replications of the deliberation-without-attention effect. *Judgment and decision making, 4*, 509.

Cunningham, S., Scerbo, M. W., & Freeman, F. G. (2000). The electrocortical correlates of daydreaming during vigilance tasks. *Journal of Mental Imagery, 24*, 61-72.

Damian, R. I., & Sherman, J. W. (2013). A process-dissociation examination of the cognitive processes underlying unconscious thought. *Journal of Experimental Social Psychology, 49*(2), 228-237. https://doi.org/10.1016/j.jesp.2012.10.018

Dijksterhuis, A. (2004). Think different: the merits of unconscious thought in preference development and decision making. *Journal of Personality and Social Psychology, 87*, 586-598. https://doi.org/10.1037/0022-3514.87.5.586

Dijksterhuis, A., Bos, M. W., Nordgren, L. F., & van Baaren, R. B. (2006). On making the right choice: the deliberation-without-attention effect. *Science, 311*, 1005-1007. https://doi.org/10.1126/science.1121629

Dijksterhuis, A., & Nordgren, L. F. (2006). A theory of unconscious thought. *Perspectives on Psychological Science, 1*, 95-109. https://doi.org/10.1111/j.1745-6916.2006.00007.x

Dijkstra, K. A., van der Pligt, J., & van Kleef, G. A. (2013). Deliberation versus intuition: Decomposing the role of expertise in judgment and decision making. *Journal of Behavioral Decision Making, 26*(3), 285-294. https://doi.org/10.1002/bdm.1759

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences [journal article]. *Behavior Research Methods, 39*, 175-191. https://doi.org/10.3758/bf03193146

Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., & Macrae, C. N. (2007). Wandering minds: The default network and stimulus-independent thought. *Science, 315*, 393-395. https://doi.org/10.1126/science.1131295

McMahon, K., Sparrow, B., Chatman, L., & Riddle, T. (2011). Driven to distraction: the impact of distracter type on unconscious decision making. *Social Cognition, 29*, 683-698. https://doi.org/10.1521/soco.2011.29.6.683

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

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review, 63*, 81-97. https://doi.org/10.1037/h0043158

Mooneyham, B. W., & Schooler, J. W. (2013). The costs and benefits of mind-wandering: A review. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 67*, 11-18. https://doi.org/10.1037/a0031569

Mrazek, M. D., Smallwood, J., Franklin, M. S., Chin, J. M., Baird, B., & Schooler, J. W. (2012). The role of mind-wandering in measurements of general aptitude. *Journal of Experimental Psychology: General, 141*(4), 788-798. https://doi.org/10.1037/a0027968

Newell, B. R., & Rakow, T. (2011). Revising beliefs about the merit of unconscious thought: Evidence in favor of the null hypothesis. *Social Cognition, 29*, 711-726. https://doi.org/10.1521/soco.2011.29.6.711

Nieuwenstein, M. R., Wierenga, T., Morey, R. D., Wicherts, J. M., Blom, T. N., Wagenmakers, E.-J., & van Rijn, H. (2015). On making the right choice: A meta-analysis and large-scale replication attempt of the unconscious thought advantage. *Judgment and Decision Making, 10*, 1-17.

Nørretranders, T. (1998). *The user illusion: cutting consciousness down to size* (Vol. 187). Viking.

Payne, J. W., Samper, A., Bettman, J. R., & Luce, M. F. (2008). Boundary conditions on unconscious thought in complex decision making. *Psychological Science, 19*, 1118-1123. https://doi.org/10.1111/j.1467-9280.2008.02212.x

Rey, A., Goldstein, R. M., & Perruchet, P. (2009). Does unconscious thought improve complex decision making? [journal article]. *Psychological Research PRPF, 73*, 372-379. https://doi.org/10.1007/s00426-008-0156-4

Robison, M. K., Miller, A. L., & Unsworth, N. (2019). Examining the effects of probe frequency, response options, and framing within the thought-probe method [journal article]. *Behavior Research Methods, 51*, 398-408. https://doi.org/10.3758/s13428-019-01212-6

Rummel, J., & Boywitt, C. D. (2014). Controlling the stream of thought: Working memory capacity predicts adjustment of mind-wandering to situational demands. *Psychonomic Bulletin & Review, 21*, 1309-1315. https://doi.org/10.3758/s13423-013-0580-3

Sayette, M. A., Reichle, E. D., & Schooler, J. W. (2009). Lost in the Sauce:The Effects of Alcohol on Mind Wandering. *Psychological Science, 20*, 747-752. https://doi.org/10.1111/j.1467-9280.2009.02351.x

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). MIT Press.

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

Seli, P., Kane, M. J., Smallwood, J., Schacter, D. L., Maillet, D., Schooler, J. W., & Smilek, D. (2018). Mind-wandering as a natural kind: A family-resemblances view. *Trends in Cognitive Sciences, 22*, 479-490. https://doi.org/10.1016/j.tics.2018.03.010

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, 4-7.*

Sio, U. N., & Ormerod, T. C. (2009). Does incubation enhance problem solving? A meta-analytic review. *Psychological Bulletin, 135*, 94-120. https://doi.org/10.1037/a0014212

Smallwood, J., & Andrews-Hanna, J. (2013). Not all minds that wander are lost: the importance of a balanced perspective on the mind-wandering state [Hypothesis and Theory]. *Frontiers in Psychology, 4*. https://doi.org/10.3389/fpsyg.2013.00441

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

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

Steindorf, L., Hammerton, H., & Rummel, J. (under review). Mind wandering outside the box – About the role of off-task thoughts and their assessment during creative incubation.

Steindorf, L., & Rummel, J. (2017). "I should not forget the apples!" - Mind-wandering episodes used as opportunities for rehearsal in an interrupted recall paradigm. *Applied Cognitive Psychology, 31*, 424-430. https://doi.org/10.1002/acp.3328

Steindorf, L., & Rummel, J. (2020). Do your eyes give you away? A validation study of eye-movement measures used as indicators for mindless reading. *Behavior Research Methods, 52,* 162-176. https://doi.org/10.3758/s13428-019-01214-4

Strick, M., Dijksterhuis, A., Bos, M. W., Sjoerdsma, A., & van Baaren, R. B. (2011). A meta-analysis on unconscious thought effects. *Social Cognition, 29*, 738-762.

Waroquier, L., Abadie, M., Klein, O., & Cleeremans, A. (2014). The effect of the cognitive demands of the distraction task on unconscious thought. *Behavioral and Brain Sciences, 37*, 44-45. https://doi.org/10.1017/S0140525X13000897

Waroquier, L., Marchiori, D., Klein, O., & Cleeremans, A. (2010). Is it better to think unconsciously or to trust your first impression? A reassessment of Unconscious Thought Theory. *Social Psychological and Personality Science, 1*, 8.

Watkins, E. R. (2008). Constructive and unconstructive repetitive thought. *Psychological Bulletin, 134*, 163-206. https://doi.org/10.1037/0033-2909.134.2.163

Weinstein, Y. (2018). Mind-wandering, how do I measure thee with probes? Let me count the ways. *Behavior Research Methods, 50*, 642-661. https://doi.org/10.3758/s13428-017-0891-9

Weinstein, Y., De Lima, H. J., & van der Zee, T. (2018). Are you mind-wandering, or is your mind on task? The effect of probe framing on mind-wandering reports. *Psychonomic Bulletin & Review, 25*, 754-760. https://doi.org/10.3758/s13423-017-1322-8

Wiemers, E. A., & Redick, T. S. (2019). The influence of thought probes on performance: Does the mind wander more if you ask it? *Psychonomic Bulletin & Review, 26*(1), 367-373. https://doi.org/10.3758/s13423-018-1529-3

Zeigarnik, B. (1927). Über das Behalten von erledigten und unerledigten Handlungen. *Psychologische Forschung, 9*, 1-85.

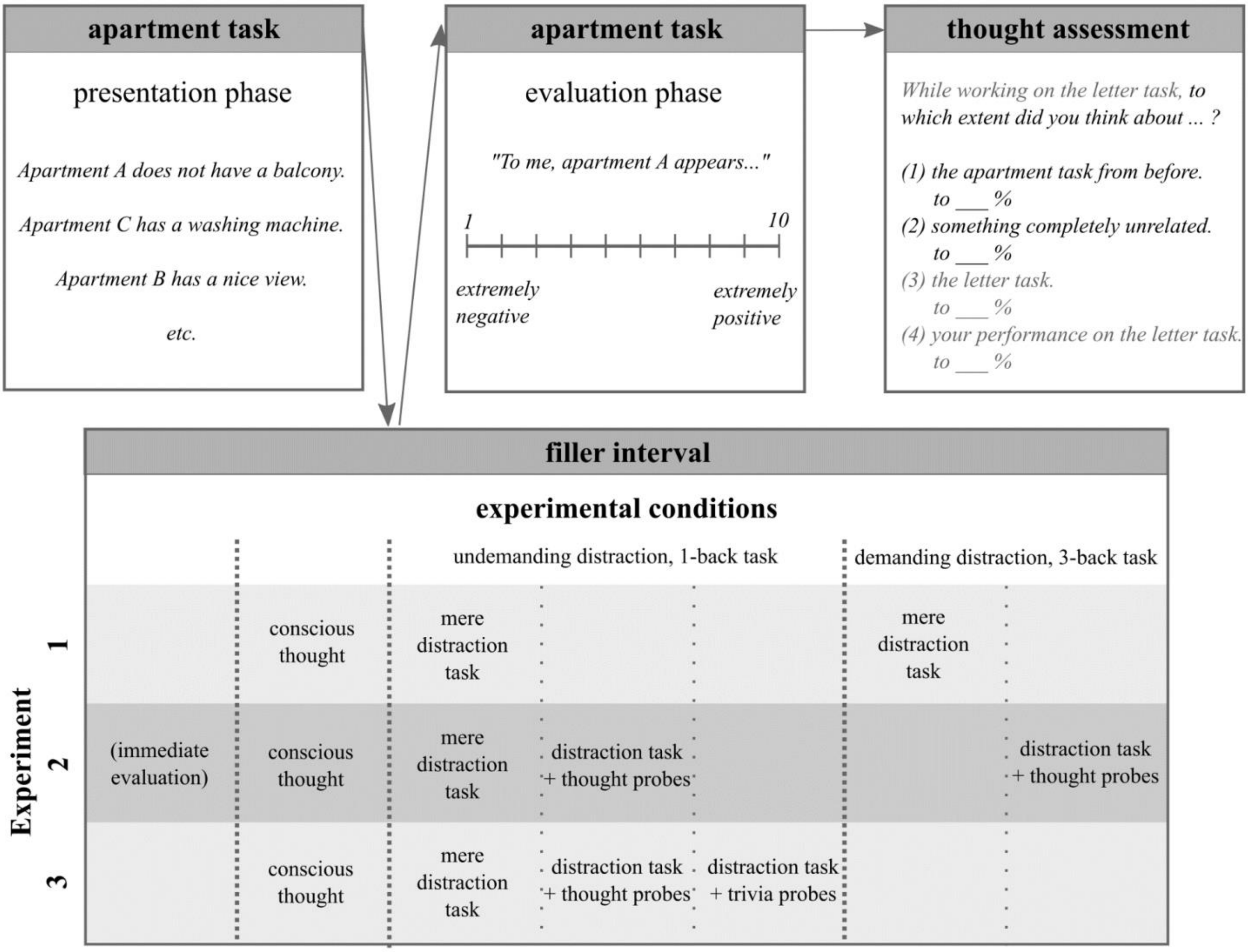


*Figure 1.* Experimental procedure. All participants worked on the apartment task's presentation phase and evaluation phase, with both phases being separated by a filler interval. The filler interval activity varied between experimental conditions and experiments. After evaluating the apartments, participants filled in a retrospective thought-assessment questionnaire, which related to their thoughts during the filler interval. Participants in the immediate evaluation condition (Experiment 2) made an exception to this general procedure, as they worked on both apartment task phases one after the other without a filler interval.

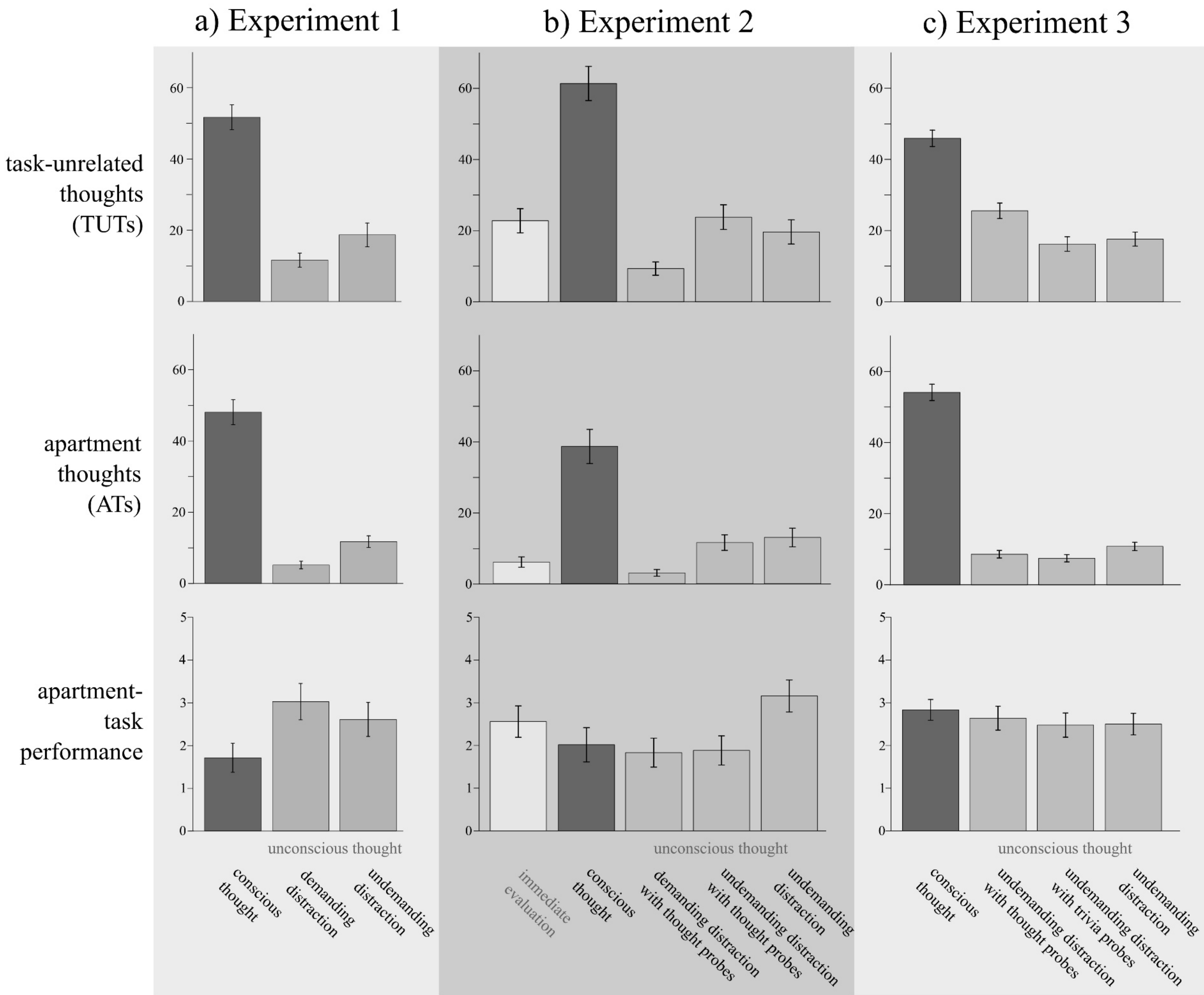
a) Experiment 1
b) Experiment 2
c) Experiment 3
task-unrelated thoughts (TUTs)
apartment thoughts (ATs)
apartment-task performance
unconscious thought
conscious thought
demanding distraction
undemanding distraction
immediate evaluation
demanding distraction with thought probes
undemanding distraction with thought probes
undemanding distraction with trivia probes

*Figure 2*. Descriptive data for the main analyses. Columns represent the respective experiment, rows the respective variable. The bars' colors stand for the thought-mode manipulations employed, with darker grey representing conscious-thought manipulations and lighter grey unconscious-thought manipulations. The immediate-evaluation condition of Experiment 2 is displayed in white. Patterns of results are explained and discussed within the running text. Error bars represent standard errors of the means.

**Appendix A4 – Manuscript 4**

Note:

This is a manuscript that is currently submitted for publication (first revision, draft version 2.5, 11th of August 2020). This version has not been peer reviewed. Please do not copy or cite without author's permission.

Steindorf, L., Hammerton, H. A., & Rummel, J. (2020). *Mind wandering outside the box – About the role of off-task thoughts and their assessment during creative incubation.*

# Mind wandering outside the box – About the role of off-task thoughts and their assessment during creative incubation

Lena Steindorf, Holly A. Hammerton, & Jan Rummel

Heidelberg University

Authors' Note

Lena Steindorf, Holly Alena Hammerton, & Jan Rummel, Department of Psychology, Heidelberg University, Heidelberg, Germany. The present research was supported by a grant from the German research foundation (DFG) to the third author (grant no: RU1996/1-1). Parts of this research were employed for the second author's unpublished bachelor thesis and presented at the 59th Annual Meeting of the Psychonomic Society, New Orleans, 2018.

Correspondence concerning this article should be addressed to Lena Steindorf, Department of Psychology, Heidelberg University, Hauptstrasse 47-51, 69117 Heidelberg, Germany. E-mail: lena.steindorf@psychologie.uni-heidelberg.de

## Abstract

The present study was designed to conceptually replicate and to further test previous findings that have shown a beneficial influence of mind wandering during incubation phases on post-incubation creative problem solving. Additionally, online thought probes and the effects their occurrence might have on incubation thought-processes were investigated. Participants worked on verbal and figural creativity tasks. In one condition, their thoughts were probed during an incubation interval, possibly interrupting and/or making aware creative thought-processes. Participants in this condition retrospectively reported fewer creativity-task-related thoughts compared to two other incubation conditions, that is, one without interruption and one interrupted by trivia questions. Creativity-task performance did not differ between these three incubation conditions and all three incubation conditions achieved a similar performance as a no-incubation control condition. These results add to the ongoing discussion regarding the relationship between mind wandering and creativity by challenging the idea of mind-wandering states contributing to a creative-incubation process.

*Keywords*: mind wandering, task-unrelated thought, creativity, incubation

Mind wandering outside the box – About the role of off-task thoughts and their assessment during creative incubation

It almost sounds self-evident: Creative thinkers need time and space to let their minds wander and to let their thoughts run freely. Tales of influential scientific thinkers solving problems following periods of distracted thought seem to have been told since the blossoming of scientific interest (e.g., Archimedes' "Eureka!" moment). However, such moments of sudden insight do not exclusively happen to famous scientists or artists. We have probably all made the experience that focusing too intently on a problem's solution does not stimulate inspiration. Sometimes, we *get stuck* and need to step away from the problem for a while. Then, out of the blue, while doing the dishes for example, the solution pops into our minds. Simple distraction tasks, such as household chores, seem to allow time for creative incubation, thereby fostering creative insight. Recent research has discussed the beneficial influence of *mind wandering* during incubation tasks on post-incubation creative performance (Baird et al., 2012; but see Smeekens & Kane, 2016). In the present study, we aimed to add to this discussion by conceptually replicating and extending previous findings. To this end, we tested for a causal link between incubation mind-wandering and creative performance. Moreover, we took a closer look at the role that different types of mind-wandering assessment might play during incubation intervals.

Many daily tasks require focused attention for successful task-goal achievement. However, it is a ubiquitous phenomenon that the mind wanders away from the task at hand towards feelings or needs, unsolved problems, personal goals or plans, and other such things (Smallwood & Andrews-Hanna, 2013). When taking an important, difficult exam, or while driving in heavy traffic, performance decrements due to off-task thoughts can come at high costs and even endanger the mind-wandering individual. Given such far-reaching consequences, it comes as no surprise that off-task thoughts have been extensively studied for their negative effects on performance in a wide range of task domains. Increased mind wandering is associated with, for instance, poorer reading comprehension (Feng, D'Mello, & Graesser, 2013; McVay & Kane, 2012; Steindorf & Rummel, 2020), lower working memory capacity (Mrazek et al., 2012; Rummel & Boywitt, 2014), lower academic achievement (Kane et al., 2017; Seli, Wammes, Risko, & Smilek, 2016; Wammes & Smilek, 2017), and poorer driving performance (Albert et al., 2018; Lin et al., 2016; Yanko & Spalek, 2014).

Since off-task thinking is found to be a pervasive phenomenon (Kane et al., 2007), it comes as good news to us frequent mind wanderers that there are benefits to match the aforementioned costs. Recent findings support the idea of mind wandering as a potentially functional practice in fields such as future and autobiographical planning (Baird, Smallwood, & Schooler, 2011; D'Argembeau, Renaud, & Van der Linden, 2011; Steindorf & Rummel, 2017), as well as in social problem solving and socio-emotional adjustment (Poerio, Totterdell, Emerson, & Miles, 2016; Ruby, Smallwood, Sackur, & Singer, 2013). Mentally travelling through time during mind-wandering episodes might also allow us to connect and compare our past and imagined future selves to the person we are today (Smallwood & Andrews-Hanna, 2013), thus providing us with a sense of self-identity, meaning, and continuity across time (Prebble, Addis, & Tippett, 2013).

In the present work, we focus on yet another construct that has been found to benefit from mind-wandering processes, namely *creative problem solving*. In the field of creativity, the question of how people produce ideas that can be defined as novel and useful (Diedrich, Benedek, Jauk, & Neubauer, 2015) is of central interest. Does creative cognition involve controlled and structured processes, or does it rely on more spontaneous and maybe even unconscious thoughts, which, while not exclusive to the drifting mind, might increasingly arise during mind-wandering episodes? Previous research suggests that both types of thought processes contribute to the production of creative solutions[1] and that the time course of creative cognition might reveal how these processes interact: In a recent review, Benedek and Jauk (2018) differentiate between a *short-term* and a *long-term perspective* of creative problem solving. The short-term perspective considers the active idea generation phase, that is, getting to know the problem at hand and actively trying to solve it, involving both controlled and spontaneous processes. According to the *controlled-attention theory of creative cognition* (Beaty, Silvia, Nusbaum, Jauk, & Benedek, 2014), this phase requires the engagement of top-down processing, which involves executive functions. The long-term perspective, however, allows for a prolonged, intermittent processing of the creative problem, following the acknowledgement that it cannot be solved right away. According to the *explicit-implicit interaction model* of creative thinking (Hélie & Sun, 2010), such an incubation phase involves unconscious and implicit associative processes, rather than conscious, explicit, and rule-governed processes.

Recent research connecting mind wandering to creativity fits well into this theoretical distinction between a short-term and long-term perspective of creative problem solving and the cognitive processes involved. Hao, Wu, Runco, and Pina (2015) found that mind wandering during active engagement in creative thought (short-term perspective, see Benedek & Jauk, 2018) is disruptive to creative performance. The authors assumed that controlled processes, which are necessary for active idea generation, are impaired by wandering

[1] Recent neuroimaging studies report the association of both default and controlled network regions with creative cognition (Jung, Mead, Carrasco, & Flores, 2013), with results from Beaty, Benedek, Kaufman, and Silvia (2015) even supporting the idea of a dynamic cooperation between both networks. The so-called default-mode network subsumes brain regions that are active when the brain is at rest and has previously been connected to mind wandering (Buckner, Andrews-Hanna, & Schacter, 2008; Christoff, Gordon, Smallwood, Smith, & Schooler, 2009).

thoughts. However, many other recent studies report a positive relationship between mind wandering and creative performance: Professional writers and physicists describe a noteworthy proportion of their most significant ideas as being formed during spontaneous mind-wandering episodes (Gable, Hopper, & Schooler, 2019). Furthermore, an increased tendency towards mind wandering was found to coincide with more *aha-experiences*, which can be described as the solving of creativity tasks due to sudden insights (Zedelius & Schooler, 2015). Similarly, mind wandering was found to be related to improved performance on commonly used tests of divergent and convergent thinking (Leszczynski et al., 2017; Preiss, Cosmelli, Grau, & Ortiz, 2016).

These results connecting mind wandering to improved creative performance can easily be integrated into the long-term perspective of creative processing (Benedek & Jauk, 2018) by assuming that “people spend more of their daily lives engaged in an incubation-like state than they probably realize” (Sawyer, 2011, p.146), and that mind-wandering individuals seem to allow more time for incubation and/or seem to make better use of this incubation time. Especially for long incubation times (e.g. a creative problem being on one’s mind for a week), mind-wandering episodes might allow for prolonged, possibly unconscious and implicit processing of the creative problem. For short incubation periods, the short-term and long-term perspectives of creative processing might overlap, and conscious, rule-governed processes might still play a role (see below).

One should keep in mind, however, that most of the just reported findings are correlational in nature. That is, they imply that people who experience creative thoughts more often also tend to mind wander more frequently or that when people experience creative thoughts, they often do so after episodes of mind wandering. Direct causal evidence for incubation mind-wandering benefitting creative problem solving is rare. In one such study, Baird et al. (2012) found that compared with a demanding task, a period of rest, or no break at all, engaging in an undemanding task during an incubation interval lead to significant improvements in performance on previously encountered versions of the Unusual Uses Task (Torrance, 1968). Furthermore, participants working on the undemanding filler task reported increased levels of off-task thoughts in a retrospective questionnaire. That is, engaging in an undemanding task during incubation was associated with both increased mind wandering and increased incubation effects on creative performance. This double association is well in line with a meta-analytic review suggesting that low-cognitive-load incubation tasks offer the greatest problem-solving benefits (Sio & Ormerod, 2009) and the finding that mind wandering is more frequent during such low-load tasks (Mason et al., 2007; Rummel & Boywitt, 2014; Smallwood, Nind, & O'Connor, 2009). It is worthwhile mentioning that Baird and colleagues (2012) did not find more thoughts explicitly related to the creativity task during the incubation period, suggesting that it is a general increase in unconscious, associative processes rather than a conscious consideration of the creativity problem during incubation mind-wandering that fosters creative performance. Interestingly, however, Smeekens and Kane (2016) did not replicate Baird and colleagues' (2012) results. In their methodologically very similar study[2], incubation mind-wandering, which was assessed via online-thought probes, did not predict post-incubation creative performance, calling the claims made by Baird and colleagues into question. Therefore, it remained unclear whether off-task thoughts during undemanding incubation tasks directly foster creative performance.

However, there was one striking distinction between the opposing studies that might account for the discrepancy between results: Whereas Baird and colleagues assessed mind wandering with retrospective questionnaires, Smeekens and Kane chose to use thought probes to assess mind wandering *online* during incubation. Consequently, the employed mind-

[2] Especially in Experiment 3, Smeekens and Kane (2016) employed similar materials to those employed in the study by Baird et al. (2012).

wandering assessment tools might have caused divergent results. Indeed, depending on the choice of measurement method, results of mind-wandering studies which applied thought probes have been found to fluctuate to an alarming level (Seli, Carriere, Levene, & Smilek, 2013; Weinstein, 2018; Weinstein, De Lima, & van der Zee, 2018, but see Schubert, Frischkorn, & Rummel, 2019). Differences in probe and instruction characteristics such as the probing frequency, the number of response options, or the framing of attentional sates can lead to changes in mind wandering and/or to increased proneness to response biases, and might affect task performance (Robison, Miller, & Unsworth, 2019). Given that even within a single type of mind-wandering assessment (online thought probes), changing probe characteristics can lead to considerable differences in results, it seems reasonable to assume that studies differing in assessment types (thought probes versus retrospective questionnaires) might produce similar or even larger discrepancies.

Concerning the methodological differences between the Baird et al. (2012) and the Smeekens and Kane (2016) studies, we assumed that in the latter online thought probing may have had disruptive effects on creative processes during incubation. This may be the reason why no incubation mind-wandering benefits were found. Indeed, creative incubation seems to be dependent upon implicit, associative processes, as are found during mind wandering (Hao et al., 2014; Hélie & Sun, 2010). Therefore, it might be possible that by continually disrupting the ongoing incubation task, the thought probes also disrupted the wandering mind's natural, associative dynamics (Smallwood & Schooler, 2015) that might have otherwise benefitted the creative incubation process. More specifically, two, not mutually exclusive, probe-related disruptions of creative thoughts are possible: First, the thought probes' merely interrupting nature might have cut off implicit, associative and possibly creative trains of off-task thought (*interruption hypothesis*). Second, thought probes might even go beyond mere interruption by alerting the participants of the key dependent measure of the experiment, namely their state of thought. In doing so, probes might have made participants more aware and more cautious of their mind-wandering behavior, possibly changing the experience itself (*awareness hypothesis*). Hence, interruption as well as awareness effects might account for the differences in results between the studies by Baird et al. (2012) and Smeekens and Kane (2016). This reasoning is based on Baird and colleagues' findings, which support the idea of unconscious, associative mind-wandering processes benefitting creative incubation, but also holds for conscious, rule-governed processes (i.e. creativity-directed mind wandering), which are crucial for the short-term perspective of creative processing (Benedek & Jauk, 2018). Such processes might also be interrupted or made aware due to thought probing.

## The present Study

The present work's goal was to test for deleterious effects of thought probes, and to thereby resolve inconsistencies between findings by Baird et al. (2012) and Smeekens and Kane (2016) to get a clearer picture of the relationship between incubation mind-wandering and creative performance. Our study qualifies as a conceptual replication because we realized a combination of both studies' methodologies, which we considered most appropriate to test our hypotheses. The inclusion of a figural creativity test (besides a verbal one) to expand the informative value of creativity results constitutes the most important methodological difference. We also made some other changes, which mostly relate to the aforementioned creativity-test expansion such as extending the amount of time participants worked on the creativity test (simply because there were more tests) or scoring creativity slightly differently (so that we could use a coherent scoring for both verbal and figural tests). Eventually, we aimed at integrating both studies' designs into one single experiment. We asked participants to work on the same divergent creativity problems (in our case verbal as well as figural problems) twice[3],

[3] We only employed repeated-exposure problems, as Baird et al. (2012) found no post-incubation benefits for new-exposure problems.

with the first and second assessment being separated by different types of incubation intervals in three experimental incubation conditions: In the first incubation condition, participants worked on an easy cognitive task (*filler task*, 0-back)[4] without any probes interrupting the task. For these participants, we expected creative incubation benefits as observed by Baird et al. (2012). In the second incubation condition, participants worked on the same task, but were interrupted by thought probes from time to time. For these participants, due to the probing procedure, we expected no creative incubation benefits, as was the case in Smeekens and Kane (2016). In the third incubation condition, participants also worked on the filler task and were interrupted from time to time. However, in this condition, we interrupted them with *trivia probes*, which had the same format as thought probes (see Method section) but consisted of general knowledge questions. We included this condition as a test for the interruption and awareness hypotheses. If mere interruption was sufficient to evoke the proposed detrimental probing effect, trivia probes should prompt such an interruption, leading to a cut-off of implicit, associative processes, thus hindering creative incubation benefits from appearing. If, however, thought awareness was a necessary criterion for the proposed detrimental nature of probes, trivia questions should not have an effect on mind wandering and/or creativity. Thus, we expanded the traditional thought-probe paradigm by adding a trivia-probe condition to gain additional insights into the cognitive processes elicited by probes. Finally, in a fourth, no-incubation condition, participants worked on both creativity assessments without any incubation time in between and worked on the same task used in the incubation condition to fill the incubation period after the second creativity assessment. This no-incubation condition thus

[4] We chose the 0-back task because Baird et al. (2012) found incubation-benefits only for simple filler tasks, which allow the mind to wander. In their study, participants working on a 0-back filler task showed improvement on the second creativity assessment whilst participants working on a more demanding 3-back task did not.

depicts a baseline condition for incubation effects, whereas the no-probe-incubation condition depicts a baseline condition for effects of thought probing on mind-wandering processes.

As a mind-wandering measure that was congruent for all conditions, every participant filled out a retrospective questionnaire concerning on-task and, more importantly, off-task thought experiences after finishing the filler task. We differentiated between general off-task thought and creative off-task thought to be able to test for influences of both kinds of thoughts on creative incubation: Creative mind wandering might be more target-oriented and might directly focus on solutions to the previously presented creativity problems. Baird et al. (2012), however, stated that such thoughts should not benefit creative incubation. Rather, implicit associative processes (i.e. general mind wandering) should lead to an increase in post-incubation creativity. One could argue that rather short incubation phases such as those employed by Baird and colleagues as well as Smeekens and Kane (2016) lie in the interstice of the above mentioned short-term and long-term perspectives of creative processing (Benedek & Jauk, 2018). Thus, either controlled processes (short term) or possibly unconscious and implicit processes (long term) might benefit creativity. The first might be represented by creative mind wandering and the latter by general mind wandering, which is why we were interested in both types of off-task thought.

In the following Method and Results sections, we report how we determined our sample size and all data exclusions, manipulations, and measures in the study (Simmons, Nelson, & Simonsohn, 2012). To follow the recommendations of Seli et al. (2018), we conceptualized mind wandering as task-unrelated thought and explained the concept to our participants accordingly. The study reported in this article was preregistered on June 01, 2017 (https://osf.io/ghwjm/; doi: 10.17605/OSF.IO/GHWJM) and, if not indicated otherwise, we followed the research and analysis protocol as stated in the preregistration. Our data is available under the same URL. Importantly, to be better able to evaluate Null effects, other than preregistered, we relied on a Bayesian approach for the present results' interpretation.

## Method

### Participants and Design

In line with a preregistered power analysis, we recruited 108 students at Heidelberg University, Germany, which were tested in groups no larger than six.[5] Data of one participant were lost due to technical errors. Another participant's data had to be excluded due to an experimenter mistake. Yet another five participants were excluded because they miscalculated their self-reported mind-wandering scores (see below), resulting in $N = 101$ ($M_{age} = 22.55$, $SD_{age} = 2.57$; 79 female, see Table 1 for the distribution across the experimental conditions) for all further analyses.

The *incubation condition* (thought-probe incubation, trivia-probe incubation, no-probe incubation, no incubation) constituted our between-participants factor; the *time of creativity assessment* (before incubation, after incubation) our repeated measures factor. The order of creativity tasks (figural, verbal) was held constant for the two measurement points but was counterbalanced across participants.

### Materials

**Verbal creativity assessment.** Participants worked on two paper-and-pencil versions (3 min per version) of the Unusual Uses Tasks (UUT) that were based on tasks from the Torrance Tests of Creative Thinking (TTCT, Torrance, 1968). Instructions were adapted from the original TTCT and translated into German by a native speaker. Participants were asked to write down as many unusual and interesting uses for a tin can (or a cardboard box in the

---

[5] As evident from the results of the Bayesian analyses, our predetermined sample size seemed large enough to allow for the interpretation of all effects relevant for our main hypotheses, thus rendering sequential sampling unnecessary.

second version) as possible within 50 response fields. They were told not to limit themselves to one size of object, to use as many objects as they wished and to not limit themselves to uses, they had already seen but instead to think of as many new ones as possible.

**Figural creativity assessment.** We adapted two versions of a figural sketching task from the TTCT. Instructions were adapted and translated into German by a native speaker. Participants were asked to sketch as many pictures with as much detail as possible within and around 42 given squares (or circles in the second version) within a three-minute time frame, and to thereby try and think of ideas nobody else could come up with. They were informed that the squares (circles) were supposed to be the center part of the picture and that they could add as many details as they liked to complete a picture. Additionally, participants were asked to add a title for each picture.

**Filler task.** Participants worked on the 0-back version of the *n*-back task. Including instructions, the filler task lasted approximately 12 min, which is comparable to the incubation time in the Baird et al. (2012) study. In the 0-back task, single letters were successively presented in the center of the screen for 500 ms each with an inter-stimulus interval of 3000 ms. Participants were asked to press a green-labeled key if the presented letter was the letter X (target trial) and to press a red-labeled key for all other letters (non-target trials). They performed one block of the 0-back task consisting of 96 non-target and 32 target trials. For all three incubation conditions (thought-probe incubation, trivia-probe incubation, no-probe incubation), the filler task was employed during the incubation period. For the no-incubation condition, the filler task was employed after the second creativity assessment.

**Online probes.** During the filler task, participants in the thought-probe-incubation and no-incubation conditions were presented with thought probes, and participants in the trivia-probe-incubation condition were presented with trivia probes after each sequence of 16 0-

back trials, resulting in eight probes in total. Both probe types had the same two-step format. First, participants were asked to categorize their current thoughts as being either task-related or task-unrelated (thought probe), or to respond with either "yes" or "no" to a random knowledge question (trivia probe, e.g., "Is the Eiffel tower higher than 300 meters?"), keeping the amount of possible responses equal for both groups. Second, as a response to an open-ended question, participants typed in a brief description of their current thoughts (thought probe), or provided a best guess concerning the exact answer to the knowledge question (trivia probe, e.g., "How high exactly?").

**Retrospective mind-wandering assessment.** To retrospectively assess the amount of mind wandering during the filler task, we asked all participants to categorize the entirety of thoughts they had experienced whilst working on this task into three categories using percentage scores: thoughts about the filler task, thoughts about the creativity problems from before, or thoughts about something completely unrelated. We considered the latter two categories as off-task thoughts, as these kinds of thoughts were unrelated to the filler task. All three percentage scores had to add up to 100, otherwise participants were excluded from the data analysis (see above).

**Procedure.**

The procedure is displayed in Figure 1. Participants signed a consent form and provided demographic information before receiving instructions for the creativity assessment. They then worked on both verbal and both figural creativity tasks in a counterbalanced order (verbal first versus figural first). Excluding instruction time, they worked on these tasks for 12 minutes (3 min per task). Afterwards, participants in the three conditions with an incubation interval between the first and second creativity assessment (thought-probe incubation, trivia-probe incubation, no-probe incubation) were told that they would stop working on the creativity tasks for now but would return to them later. They received instructions for the filler task and the respective probing procedure (where applicable), including a detailed definition of mind wandering (where applicable) as well as examples. They then worked on the filler task, completed the retrospective mind-wandering questionnaire (on which mind wandering was also defined and described), and then started the second creativity assessment. The creativity problems were presented in the same order as during the first assessment. After completing the first creativity assessment, participants in the no-incubation condition immediately started the second creativity assessment, following the information that they would now once again work on the creativity problems in the same order as before. After having finished both creativity assessments, no-incubation participants received instructions for the filler task and the thought-probing procedure, worked on this task, and completed the retrospective mind-wandering questionnaire. Before the second creativity assessment, all participants were told that they could re-use or adapt their answers from the first creativity assessment, but that coming up with completely new and unusual answers to the creativity problems would result in higher test scores. Finally, all participants filled in a trait-based mind-wandering questionnaire[6] and were then debriefed and dismissed.

## Results

We chose to employ Bayesian analyses in addition to the preregistered frequentist analyses to gain richer information and draw more informative inferences regarding the likelihood of our data to occur under the assumptions of the Null as well as the alternative hypotheses. In the following sections, we report the common information for frequentist ANOVAs and ANCOVAs, but use the additionally reported Bayes factors to draw conclusions concerning our hypotheses. To this end, for all of our analyses, we either report

[6] As a more trait-based measure of mind wandering, we adapted a scale by Carriere, Seli, and Smilek (2013). Unfortunately, these data were not recorded due to programming errors and thus will not be included in later analyses.

$BF_{10}$ as the Bayes factor for the alternative versus the Null hypothesis, or $BF_{01}$ as the Bayes factor for the Null versus the alternative hypothesis. For all reported Bayesian ANOVAs and ANCOVAs (BANOVAs and BANCOVAs), we employed Cauchy priors (Wagenmakers et al., 2018) and followed the guidelines by Jeffreys (1961) for interpretation[7]. The analyses were executed using JASP (JASP Team, 2019). The order manipulation of creativity tasks (figural first, verbal first), which was counterbalanced between participants, yielded main effects for some of the creativity measures (probably reflecting motivational decreases over time) but did not interact with condition effects. Thus, we collapsed across this factor for all reported analyses.

**Filler task performance.** For the 0-back filler task, performance was defined as the percentage of trials (target and non-target) with correct responses in relation to all trials. There were no significant differences between conditions for 0-back task performance, $F(3, 97) = 0.70$, $p = .556$, $\eta^2_p = .02$ (see Table 1 for descriptive statistics), as suggested by a one factorial ANOVA with incubation condition as between-participants factor. The BANOVA revealed substantial support for the Null hypothesis, by showing that our data were more than eight times more likely to occur under a model excluding an effect for the experimental condition, $BF_{01} = 8.62$. That is, 0-back task performance was not affected by the kind of incubation (or no incubation) employed in the respective experimental conditions.

**Mind-wandering behavior.** The amount of mind wandering measured online via forced-choice thought probes (first thought-probing step) during the filler task was defined as the percentage of probes in which participants categorized themselves as being off-task. Thought probes were employed for two conditions only: Participants' mind-wandering behavior, measured via probes, did not differ for the thought-probe incubation compared to

[7] *BF* 1-3: anecdotal support; *BF* 3-10: substantial support; *BF* 10-30: strong support; *BF* 30-100: very strong support; *BF* > 100: decisive support

the no-incubation condition, $t(48) = 0.01$, $p = .994$, $d = 002$, $BF_{01} = 3.53$ (see Table 1 for descriptive statistics), as was suggested by an independent samples *t*-test. That is, overall mind-wandering frequency, measured online during the filler task, was not influenced by whether the creativity task was interrupted or already finished during assessment.

The number of thoughts classified as creativity thoughts (that is, thoughts concerning the creativity tasks) from the open-ended thought-probe questions (second thought-probing step) was very low, altogether. For this measure, one rater evaluated the thought descriptions participants had provided. In the thought-probe incubation, as well as in the no-incubation condition, only two participants reported having experienced thoughts about the creativity task. For this reason, we refrained from running any further statistical analyses for this measure.

The amount of mind wandering assessed retrospectively after the filler task was specified using percentage scores. Overall, participants reported having been on-task for 67.24 percent ($SD = 20.15$) of the time. A one factorial (B)ANOVA with incubation condition as between-participants factor showed that the amount of off-task thoughts about the creativity problems varied between conditions $F(3, 97) = 3.60$, $p = .016$, $\eta^2_p = .10$, $BF_{10} = 2.76$. Our primary hypothesis concerned a potential reactivity effect of thought probes and its origin. To test for this effect, we used a set of Helmert contrasts. The first contrast tested the thought probe incubation condition against the no probe incubation condition and the trivia probe condition to investigate whether thought probing would systematically affect retrospective thought reports. This was the case, $F(1, 97) = 4.99$, $p = .028$, $\eta^2_p = .06$, $BF_{10} = 2.02$. The second contrast tested for differences between the no-probe-incubation condition and the trivia-probe condition and revealed that the amount of creativity thoughts did not differ between the two, $F(1, 97) = 0.22$, $p = .639$, $\eta^2_p = .01$, $BF_{01} = 3.25$. To achieve a more complete picture of the thought probe effect, we used additional post-hoc comparisons to test each incubation condition against the no-incubation condition. These comparisons revealed increased amounts of creativity thoughts in the trivia-probe incubation, $t(97) = -2.76$, $p = .034$, $BF_{10} = 5.81$, and, however with anecdotal evidence, in the no-probe incubation conditions, $t(97) = -2.21$, $p = .129$, $BF_{10} = 2.84$, but not in the thought-probe incubation condition, $t(97) = -047$, $p = .966$, $BF_{01} = 3.11$.

The amount of other, general off-task thoughts (excluding creativity thoughts) did not vary between conditions $F(3, 97) = 0.43$, $p = .733$, $\eta^2_p = .01$, $BF_{01} = 11.59$ (see Figure 2), as was suggested by the respective (B)ANOVA. That is, we did not find higher levels of general mind wandering (excluding creative thoughts) during the 0-back task for any of the incubation conditions as compared to the no-incubation condition.

**Verbal creativity.** Two independent raters evaluated the meaningfulness of all responses to both UUT versions (tin can and cardboard box), so that far-fetched, non-viable answers could be excluded from further analyses. Discrepancies in meaningfulness ratings were resolved by a third, independent rater. After excluding non-meaningful responses, we computed individual scores on three different measures of verbal creativity: *originality*, *uniqueness*, and *fluency*. The same two raters who had evaluated the meaningfulness of responses assessed the originality of responses to both UUT versions using a scale of zero (“The idea does not extend beyond the traditional use of a tin can/cardboard box.”) to three (“The idea is unexpected, inventive, and requires an extraordinary capacity for outside-the-box thinking”). The two raters’ scores were averaged for each response and summed up for each participant and creativity assessment over both UUT versions, creating two originality scores (one for the first, one for the second creativity assessment) for each participant. This subjective measure of creativity was similar to the one used by Smeekens and Kane (2016). Additionally, to produce a more objective measure of originality (see Baird et al., 2011), we calculated two uniqueness scores (one per creativity assessment) for each participant.

Responses were pooled across all participants separately for the tin can and the cardboard box versions of the UUT and each response was assigned a uniqueness value, which was computed by dividing one by the number of times the response had been given for the respective UUT version. That is, the less often a response had been given, the closer its uniqueness value was to one. Separately for the two creativity assessments, uniqueness values were summed up for each participant over both UUT versions, creating two uniqueness scores for each participant (one for the first, one for the second creativity assessment). As a third creativity measure, we calculated a fluency score for each participant and each creativity assessment, which simply represented the number of meaningful responses a participant had given to both UUT versions.

To test whether verbal creativity performance in the second assessment differed between conditions (that is, depending on whether there is an incubation phase and depending on the kind of incubation) whilst controlling for a priori differences in the first assessment, we ran frequentist and Bayesian ANCOVAs[8] with the respective verbal creativity measure for the second assessment as the dependent variable, incubation condition as a between-participants factor, and the respective verbal creativity measure for the first assessment as a covariate. Table 2 contains the descriptive statistics for all three verbal creativity measures and both assessments. We did not find any differences between conditions regarding originality in the second creativity assessment, $F(3, 96) = 0.21$, $p = .891$, $\eta^2_p = .00$, $BF_{01} = 15.46$, as was the case for uniqueness in the second creativity assessment, $F(3, 96) = 0.40$, $p = .754$, $\eta^2_p = .01$, $BF_{01} = 10.96$. Regarding creative fluency in the second assessment, we again did not find differences between experimental conditions, $F(3, 96) = 0.18$, $p = .913$, $\eta^2_p = .00$, $BF_{01} =$

[8] Other than preregistered, we decided to report (B)ANCOVAs rather than repeated-measure (B)ANOVAs, because there was no significant interaction between time of creativity assessment and incubation condition.

13.70. For all three creativity measures, performance in the first assessment was a significant covariate (each with $p < .001$ and a $BF_{10}$ with decisive support). That is, controlling for the performance in the first assessment, verbal creativity performance in the second assessment did not vary between groups. Neither incubation per se, nor a special kind of incubation was found to produce creative performance benefits.

**Figural creativity.** For figural creativity, a condition-blinded rater evaluated the meaningfulness of all responses to both versions of the figural sketching task (squares and circles), so that sketches with improper uses of the squares (circles) could be excluded for further analyses in the next step. For figural creativity, we computed individual scores on two different measures: *originality* and *fluency*. Similar to the procedure for verbal creativity, a rater assessed the originality of all sketches using a scale of zero ("The idea does not extend beyond the meaning of a square/circle.") to three ("The idea is unexpected, inventive, and requires an extraordinary capacity for outside-the-box thinking") taking into account the respective sketch and its title. Scores were summed up for each participant and creativity assessment over both versions of the figural sketching task, creating two originality scores (one for the first, one for the second creativity assessment) for each participant. The fluency score for each participant and each creativity assessment represented the number of meaningful sketches a participant had drawn for both the square and the circle version of the task. We did not compute uniqueness scores for figural creativity, because the sketches were not as objectively comparable as answers to the UUT. Because five participants did not produce meaningful sketches on all task versions, we excluded them from all further analyses for figural creativity performance, resulting in $N = 96$ for these analyses.

We again ran frequentist and Bayesian ANCOVAs with the respective figural creativity measure for the second assessment as dependent variable, incubation condition as between-participants factor, and the respective figural creativity measure for the first assessment as a covariate (see Table 2 for descriptive statistics). We did not find differences between experimental conditions for originality in the second creativity assessment, $F(3, 91) = 0.62$, $p = .606$, $\eta^2_p = .01$, $BF_{01} = 4.11$, or fluency in the second assessment, $F(3, 91) = 0.57$, $p = .639$, $\eta^2_p = .01$, $BF_{01} = 2.77$. As for verbal creativity, we conclude that neither incubation per se, nor a special kind of incubation was found to produce creative performance benefits.

**Correlational analyses.** We correlated incubation mind-wandering with post-incubation creativity performance. For the thought-probe incubation condition (similar to Smeekens & Kane, 2016), as well as for the three incubation conditions combined (thought-probe incubation, trivia-probe incubation, no-probe incubation), we obtained non-significant correlations only, all *p*s > .1 and $BF_{10} < 1$ (for (Bayesian) Pearson correlations) when correlating the mind-wandering measures (probe-caught general off-task thoughts, retrospectively reported general off-task thoughts, and retrospectively reported creativity-related thoughts) with the creativity performance measures from the second assessment (all verbal and figural measures), which is well in line with Smeekens and Kane's results.

## Discussion

Although the last couple of years have brought forth a promising amount of research supporting the idea that mind wandering could foster creative problem solving (Gable et al., 2019; Hao et al., 2014; Leszczynski et al., 2017; Preiss et al., 2016; Zedelius & Schooler, 2015), the specific functionality of off-task thoughts during short incubation intervals remained unclear. Considering the long-term perspective of creative problem solving (see Introduction and Benedek & Jauk, 2018), such incubation intervals should involve unconscious and implicit associative processes, which, while not exclusive to the wandering mind, might increasingly arise during off-task thought episodes and might benefit creativity. Considering the short-term perspective, conscious, rule-governed processes in the form of creative mind wandering might also play a role fostering creativity. However, mixed results regarding the causal relation between incubation mind-wandering and creative problem solving had been found in previous studies (Baird et al., 2012; Smeekens & Kane, 2016) and we hypothesized that these inconsistencies could have been caused by methodological differences in mind-wandering assessment.

In the present study, we had expected to find creative incubation benefits at least in the no-probe incubation condition and possibly also in the trivia-probe incubation condition (if an increased meta-awareness about one's mental states, and not only mere interruption, hampered the mind-wandering–creativity link). Such a pattern of results would have supported the assumption that mind wandering indeed facilitates creative incubation, as long as the mind is allowed to actually run freely without any disturbance. However, we found neither a post-incubation creative performance benefit for the three incubation groups (thought-probe incubation, trivia-probe incubation, no-probe incubation) in contrast to the no-incubation group, nor any benefits for special kinds of incubation. Additionally, differences in mind-wandering behavior evoked by our experimental manipulation (discussed below) were not related to creative performance in the second assessment. Neither conscious thoughts directed at the creativity problem nor general, possibly implicit and associative mind-wandering processes were associated with post-incubation creativity.

Not only were we interested in the causal link between incubation mind-wandering and post-incubation creative performance, our expansion of the traditional thought-probe paradigm was also supposed to generate new knowledge on the nature of thought probes and the (possibly detrimental) effects they might evoke. More specifically, our employed paradigm was supposed to test the interruption and the awareness hypothesis (see Introduction). Retrospectively applied mind-wandering questionnaires revealed that there were no group differences concerning general mind wandering excluding creativity thoughts. However, our new paradigm allowed us to observe group differences concerning thoughts about the creativity tasks. We found higher levels of creativity thoughts for two of our incubation groups (trivia-probe incubation and no-probe incubation) in comparison to the baseline no-incubation group. This increase could be described as a *Zeigarnik*-like effect (Zeigarnik, 1927). Interrupting the creativity assessment with an incubation phase could have led to an increased availability of and an increased mental occupation with the demands and contents of the creativity problems[9]. However, interestingly, asking participants about their thoughts led to an absence of this increase in the thought-probe incubation group. This absence is well in line with our proposed awareness hypothesis. Mere interruption proved insufficient to change participant's mind-wandering behavior in contrast to a no-probe condition. However, going beyond mere interruption by alerting the participants of the key dependent measure of the experiment, namely their state of thought, might have made participants more aware and more cautious of their wandering minds. More precisely, not only thought-probe characteristics (Seli et al., 2013; Weinstein, 2018; Weinstein et al., 2018), but also the mere occurrence of such probes seems to reactively affect creativity-task-related mind wandering through increased thought awareness.

Alternative explanations for the group differences concerning creativity thoughts may exist. For example, thought probes might enable more accurate retrospective thought reports by providing participants with initial estimates of the amount of creativity thoughts. Participants without such initial estimates might retrospectively report inflated creativity-thought levels due to demand characteristics of the questionnaire.[10] Although plausible, the literature on Zeigarnik-like effects deems our initial interpretation of the finding of increased

[9]Although plausible, our results conflict with the findings from Baird et al. (2012), because their undemanding filler task condition was characterized by high levels of general off-task thought but not of thoughts especially directed toward the creativity tasks.

[10] We thank one of our reviewers for raising this point.

creative thought for incubation participants (excluding thought-probe-incubation participants) in contrast to no-incubation participants as such a Zeigarnik-like effect more likely. Such an increase is well in line with other studies reporting a prolonged mental occupation with still active problems or tasks during incubation intervals (e.g., Bugg & Scullin, 2013; Steindorf & Rummel, 2017). In this light, it seems more likely that our creativity-thought rates reflect a prolonged mental occupation with the pending creativity task rather than occurring due to demand characteristics. Consequently, we believe the absence of the creativity-thought increase in the thought-probe-incubation condition to be due to an absence of a Zeigarnik-like effect (due to effects of thought awareness) rather than a more accurate creativity-thought estimation. However, further research is needed to support this interpretation.

A recent study (Wiemers & Redick, 2019) described thought probes as a non-reactive method for the measurement of mind wandering when embedding them into a sustained attention to response task. Our results support the notion that probes do not alter performance for current cognitive tasks. However, mere thought-probe occurrences might still alter mind-wandering contents during such tasks and, for example, hinder associative or goal-related processes from taking place. Performance in still active, unfinished or interrupted tasks, which benefit from or depend on such processes, might thereby also be affected by the occurrence of thought probes. However, this was the first time we employed an extended version of the traditional thought-probing procedure to disentangle interruption and awareness effects on different types of mind wandering. Further research is needed to shed light on the influence of thought probing on mind wandering, and we consider our extended paradigm to be a very fruitful approach towards this end. For this reason, we are currently examining reactive effects of thought probing on mind wandering processes within other problem-solving scenarios, such as the unconscious thought paradigm (Steindorf, Rummel, & Boywitt, manuscript submitted for publication).

Even though our results suggest that online thought probing alters goal-related mind wandering during incubation, these changes within participants' mental states were not related to changes in creativity. We did not find evidence for a causal link between incubation mind-wandering and post-incubation creative performance. Creativity could not be situationally increased with mind-wandering manipulations. However, this argumentation does not rule out the possibility that incubation mind-wandering might be found to be beneficial within paradigms producing reliable incubation effects. In our study, we focused on divergent creativity measures. Performance on convergent creativity tasks, for example, might benefit from incubation mind-wandering. In these tasks, people are supposed to find one particular correct solution to a problem. Specific convergent-creativity-related off-task thoughts during an incubation interval might foster post-incubation performance, because they might constitute an active search (requiring top-down control, see Colzato, Szapora, & Hommel, 2012) for the one correct solution.

A positive relationship between mind wandering and creativity seems to be most prevalently found when mind wandering is assessed on a trait level. That is, creative people might indeed tend to mind wander more, or in the other direction, frequent mind-wanderers might indeed be more creative. However, manipulating state mind wandering to increase post-incubation creativity in an artificial lab environment might not be possible. Offering people the possibility to mind wander (in our case through an easy filler task) might not be sufficient for creative incubation.

Furthermore, one could argue that the relationship between mind wandering and creativity is (partially) driven by stereotypes and self-images concerning creative and daydreaming people. For example, people who are more creative quite probably identify themselves with the stereotype regarding "the creative person", which could possibly cause them to see mind wandering as a positive process that could benefit finding inspiration, and which “belongs” to their creative identity. Subsequently, when asked to report their mind-wandering propensity, their answers might be tinted by their self-image, thus artificially increasing the amount of reported mind wandering compared to less creative individuals.

In conclusion, in the present work, employing an extended thought probing paradigm, we observed mind-wandering changes due to the mere existence of thought probes. Our aim had been to relate those changes in incubation mind-wandering to post-incubation creative performance to resolve inconsistencies between findings by Baird et al. (2012) and Smeekens and Kane (2016). Replicating only the result found by Smeekens and Kane, in-the-moment mind wandering during incubation was not associated with post-incubation creativity. However, from a daily-life and a rather trait-based perspective, we should not discard the idea of a positive relationship between creativity and mind wandering.

**References**

Albert, D. A., Ouimet, M. C., Jarret, J., Cloutier, M.-S., Paquette, M., Badeau, N., & Brown, T. G. (2018). Linking mind wandering tendency to risky driving in young male drivers. *Accident Analysis & Prevention, 111*, 125-132. doi:10.1016/j.aap.2017.11.019

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. doi: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*, 1604-1611. doi:10.1016/j.concog.2011.08.007

Beaty, R. E., Benedek, M., Kaufman, B. S., & Silvia, P. J. (2015). Default and executive network coupling supports creative idea production. *Scientific Reports, 5*, 10964. doi:10.1038/srep10964

Beaty, R. E., Silvia, P. J., Nusbaum, E. C., Jauk, E., & Benedek, M. (2014). The roles of associative and executive processes in creative cognition. *Memory & Cognition, 42*, 1186-1197. doi:10.3758/s13421-014-0428-8

Benedek, M., & Jauk, E. (2018). Spontaneous and controlled processes in creative cognition. In K. C. R. Fox & K. Christoff (Eds.), *The Oxford handbook of spontaneous thought: Mind-wandering, creativity, and dreaming*. New York, NY: Oxford University Press.

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

Bugg, J. M., & Scullin, M. K. (2013). Controlling intentions: The surprising ease of stopping after going relative to stopping after never having gone. *Psychological Science*, *24*, 2463-2471. doi: 10.1177/0956797613494850

Carriere, J. S., Seli, P., & Smilek, D. (2013). Wandering in both mind and body: individual differences in mind wandering and inattention predict fidgeting. *Canadian Journal of Experimental Psychology, 67*, 19-31. doi:10.1037/a0031438

Christoff, K., Gordon, A. M., Smallwood, J., Smith, R., & Schooler, J. W. (2009). Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proceedings of the National Academy of Sciences, 106*, 8719-8724. doi:10.1073/pnas.0900234106

Colzato, L., Szapora, A., & Hommel, B. (2012). Meditate to create: the impact of focused-attention and open-monitoring training on convergent and divergent thinking. *Frontiers in Psychology, 3,* 116. doi:10.3389/fpsyg.2012.00116

D'Argembeau, A., Renaud, O., & Van der Linden, M. (2011). Frequency, characteristics and functions of future-oriented thoughts in daily life. *Applied Cognitive Psychology, 25*, 96-103. doi:10.1002/acp.1647

Diedrich, J., Benedek, M., Jauk, E., & Neubauer, A. C. (2015). Are creative ideas novel and useful? *Psychology of Aesthetics, Creativity, and the Arts, 9*, 35-40. doi:10.1037/a0038688

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

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*. doi:10.1177/0956797618820626

Hao, N., Ku, Y., Liu, M., Hu, Y., Grabner, R. H., & Fink, A. (2014). Enhancing verbal creativity via brief interventions during an incubation interval. *Creativity Research Journal, 26*, 30-38. doi:10.1080/10400419.2014.873658

Hao, N., Wu, M., Runco, M. A., & Pina, J. (2015). More mind wandering, fewer original ideas: Be not distracted during creative idea generation. *Acta Psychologica, 161*, 110-116. doi:10.1016/j.actpsy.2015.09.001

Hélie, S., & Sun, R. (2010). Incubation, insight, and creative problem solving: a unified theory and a connectionist model. *Psychological Review, 117*, 994-1024. doi:10.1037/a0019532

Jeffreys, H. (1961). *Theory of probability*. Oxford, UK: Oxford University Press.

Jung, R. E., Mead, B. S., Carrasco, J., & Flores, R. A. (2013). The structure of creative cognition in the human brain. *Frontiers in Human Neuroscience, 7*, 330. doi:10.3389/fnhum.2013.00330

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. doi:10.1111/j.1467-9280.2007.01948.x

Kane, M. J., Smeekens, B. A., von Bastian, C. C., Lurquin, J. H., Carruth, N. P., & Miyake, A. (2017). A combined experimental and individual-differences investigation into mind wandering during a video lecture. *Journal of Experimental Psychology: General, 146*, 1649-1674. doi:10.1037/xge0000362

Leszczynski, M., Chaieb, L., Reber, T. P., Derner, M., Axmacher, N., & Fell, J. (2017). Mind wandering simultaneously prolongs reactions and promotes creative incubation. *Scientific Reports, 7*, 10197. doi:10.1038/s41598-017-10616-3

Lin, C.-T., Chuang, C.-H., Kerick, S., Mullen, T., Jung, T.-P., Ko, L.-W., . . . McDowell, K. (2016). Mind-wandering tends to occur under low perceptual demands during driving. *Scientific Reports, 6*, 21353. doi:10.1038/srep21353

Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., & Macrae, C. N. (2007). Wandering minds: the default network and stimulus-independent thought. *Science, 315*, 393-395. doi:10.1126/science.1131295

McVay, J. C., & Kane, M. J. (2012). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of Experimental Psychology: General, 141*, 302-320. doi:10.1037/a0025250

Mrazek, M. D., Smallwood, J., Franklin, M. S., Chin, J. M., Baird, B., & Schooler, J. W. (2012). The role of mind-wandering in measurements of general aptitude. *Journal of Experimental Psychology: General, 141*, 788-798. doi:10.1037/a0027968

Poerio, G. L., Totterdell, P., Emerson, L.-M., & Miles, E. (2016). Social daydreaming and adjustment: an experience-sampling study of socio-emotional adaptation during a life transition. *Frontiers in Psychology, 7,* 13. doi:10.3389/fpsyg.2016.00013

Prebble, S. C., Addis, D. R., & Tippett, L. J. (2013). Autobiographical memory and sense of self. *Psychological Bulletin, 139*, 815-840. doi:10.1037/a0030146

Preiss, D. D., Cosmelli, D., Grau, V., & Ortiz, D. (2016). Examining the influence of mind wandering and metacognition on creativity in university and vocational students. *Learning and Individual Differences, 51*, 417-426. doi:10.1016/j.lindif.2016.07.010

Robison, M. K., Miller, A. L., & Unsworth, N. (2019). Examining the effects of probe frequency, response options, and framing within the thought-probe method. *Behavior Research Methods, 51*, 398-408. doi:10.3758/s13428-019-01212-6

Ruby, F. J. M., Smallwood, J., Sackur, J., & Singer, T. (2013). Is self-generated thought a means of social problem solving? *Frontiers in Psychology, 4*, 962-962. doi:10.3389/fpsyg.2013.00962

Rummel, J., & Boywitt, C. D. (2014). Controlling the stream of thought: Working memory capacity predicts adjustment of mind-wandering to situational demands. *Psychonomic Bulletin & Review, 21*, 1309-1315. doi:10.3758/s13423-013-0580-3

Sawyer, K. (2011). The cognitive neuroscience of creativity: a critical review. *Creativity Research Journal, 23*, 137-154. doi:10.1080/10400419.2011.571191

Schubert, A. L., Frischkorn, G. T., & Rummel, J. (2019). The validity of the online thought-probing procedure of mind wandering is not threatened by variations of probe rate and probe framing. *Psychological Research*, Advanced online publication. doi:10.1007/s00426-019-01194-2

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

Seli, P., Kane, M. J., Smallwood, J., Schacter, D. L., Maillet, D., Schooler, J. W., & Smilek, D. (2018). Mind-wandering as a natural kind: A family-resemblances view. *Trends in Cognitive Science, 22*, 479-490. doi:10.1016/j.tics.2018.03.010

Seli, P., Wammes, J. D., Risko, E. F., & Smilek, D. (2016). On the relation between motivation and retention in educational contexts: the role of intentional and unintentional mind wandering. *Psychonomic Bulletin & Review, 23*, 1280-1287. doi:10.3758/s13423-015-0979-0

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, 4-7.*

Sio, U. N., & Ormerod, T. C. (2009). Does incubation enhance problem solving? A meta-analytic review. *Psychological Bulletin, 135*, 94-120. doi:10.1037/a0014212

Smallwood, J., & Andrews-Hanna, J. (2013). Not all minds that wander are lost: the importance of a balanced perspective on the mind-wandering state. *Frontiers in Psychology, 4,* 441. doi:10.3389/fpsyg.2013.00441

Smallwood, J., Nind, L., & O'Connor, R. C. (2009). When is your head at? An exploration of the factors associated with the temporal focus of the wandering mind. *Consciousness and Cognition, 18*, 118-125. doi:10.1016/j.concog.2008.11.004

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

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*, 389-415. doi:10.1037/aca0000046

Steindorf, L., & Rummel, J. (2017). "I should not forget the apples!" - Mind-wandering episodes used as opportunities for rehearsal in an interrupted recall paradigm. *Applied Cognitive Psychology, 31*, 424-430. doi:10.1002/acp.3328

Steindorf, L., & Rummel, J. (2020). Do your eyes give you away? A validation study of eye-movement measures used as indicators for mindless reading. *Behavior Research Methods, 52,* 162-176. doi:10.3758/s13428-019-01214-4

Steindorf, L., & Rummel, J., & Boywitt, C. D. (2020). *A fresh look at the Unconscious Thought Effect: Using mind-wandering measures to reveal thought processes in decision problems with high information load.* Manuscript submitted for publication.

JASP Team (2019). JASP (Version 0.11.1) [Computer software].

Torrance, E. P. (1968). *Torrance tests of creative thinking*: Princeton: NJ: Personell Press.

Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., . . . Morey, R. D. (2018). Bayesian inference for psychology. Part II: example applications with JASP. *Psychonomic Bulletin & Review, 25*, 58-76. doi:10.3758/s13423-017-1323-7

Wammes, J. D., & Smilek, D. (2017). Examining the influence of lecture format on degree of wind wandering. *Journal of Applied Research in Memory and Cognition, 6*, 174-184. doi:10.1016/j.jarmac.2017.01.015

Weinstein, Y. (2018). Mind-wandering, how do I measure thee with probes? Let me count the ways. *Behavior Research Methods, 50*, 642-661. doi:10.3758/s13428-017-0891-9

Weinstein, Y., De Lima, H. J., & van der Zee, T. (2018). Are you mind-wandering, or is your mind on task? The effect of probe framing on mind-wandering reports. *Psychonomic Bulletin & Review, 25*, 754-760. doi:10.3758/s13423-017-1322-8

Wiemers, E. A., & Redick, T. S. (2019). The influence of thought probes on performance: Does the mind wander more if you ask it? *Psychonomic Bulletin & Review, 26*, 367-373. doi:10.3758/s13423-018-1529-3

Yanko, M. R., & Spalek, T. M. (2014). Driving with the wandering mind: the effect that mind-wandering has on driving performance. *Human Factors, 56*, 260-269. doi:10.1177/0018720813495280

Zedelius, C. M., & Schooler, J. W. (2015). Mind wandering “Ahas” versus mindful reasoning: alternative routes to creative solutions. *Frontiers in Psychology, 6,* 834. doi:10.3389/fpsyg.2015.00834

Zeigarnik, B. (1927). Über das Behalten von erledigten und unerledigten Handlungen. *Psychologische Forschung, 9*, 1-85.

Table 1. *Distribution of N across the experimental conditions, as well as means (M) and standard deviations (SD) for filler task (0-back task) performance and probed mind wandering.*

| Condition | | 0-back (% correct) | | Off-task thoughts (%, thought probes) | |
|---|---|---|---|---|---|
| | *N* | *M* | *SD* | *M* | *SD* |
| thought-probe incubation | 26 | 93.00 | 4.03 | 25.48 | 19.20 |
| trivia-probe incubation | 26 | 91.50 | 6.16 | not assessed | |
| no-probe incubation | 25 | 93.16 | 3.95 | | |
| no incubation | 24 | 92.71 | 3.63 | 25.52 | 18.97 |

Table 2. *Means (M) and standard deviations (SD) for creativity performance in the first and second assessment.*

*Note.* Values for verbal creativity are based on $N$ = 101. Values for figural creativity are based on $N$ = 96.

| | verbal creativity | | | | | | figural creativity | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | originality | | uniqueness | | fluency | | originality | | fluency | |
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| | first assessment | | | | | | | | | |
| thought-probe incubation | 27.38 | 12.10 | 3.39 | 1.89 | 20.62 | 7.21 | 11.42 | 5.38 | 7.54 | 3.78 |
| trivia-probe incubation | 24.40 | 11.95 | 2.79 | 1.85 | 18.69 | 7.72 | 12.09 | 5.79 | 8.44 | 3.67 |
| no-probe incubation | 29.50 | 11.81 | 3.14 | 1.50 | 22.64 | 7.72 | 10.72 | 7.58 | 7.20 | 5.38 |
| no incubation | 28.73 | 10.04 | 3.28 | 1.88 | 21.75 | 5.54 | 13.64 | 7.63 | 9.46 | 4.91 |
| | second assessment | | | | | | | | | |
| thought-probe incubation | 26.29 | 12.20 | 4.26 | 2.81 | 19.00 | 7.87 | 14.12 | 7.27 | 9.08 | 3.93 |
| trivia-probe incubation | 24.38 | 11.05 | 3.53 | 2.67 | 18.00 | 6.91 | 15.30 | 5.72 | 10.09 | 4.03 |
| no-probe incubation | 26.12 | 10.70 | 3.64 | 2.38 | 19.88 | 8.00 | 12.44 | 9.21 | 8.28 | 6.07 |
| no incubation | 25.79 | 9.26 | 3.60 | 1.85 | 19.08 | 6.83 | 16.73 | 7.69 | 11.27 | 4.74 |

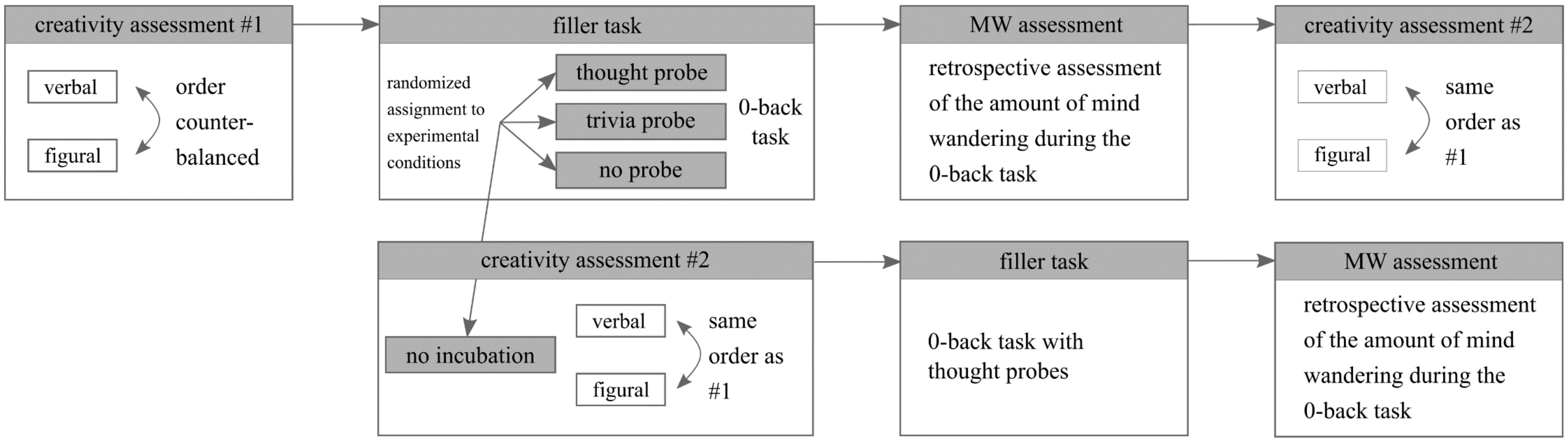


*Figure 1*. Experimental procedure. Participants in the conditions including an incubation interval (thought-probe incubation, trivia-probe incubation, no-probe incubation) worked on the 0-back task between the two creativity assessments, as is displayed in the upper half of the figure. Participants in the no-incubation condition worked on the 0-back task after having finished both creativity assessments, as is displayed in the lower half of the figure.

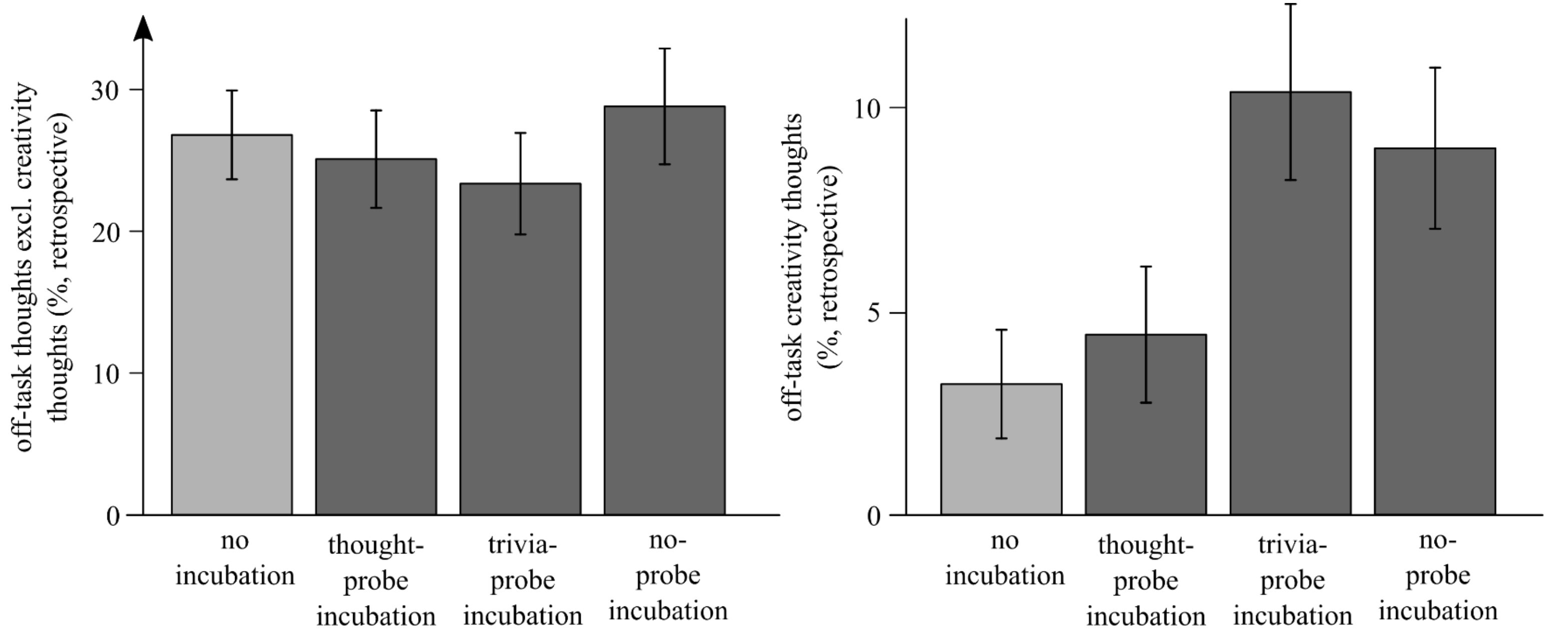


*Figure 2.* Mean values for retrospectively reported mind wandering. The figure on the left depicts general off-task thoughts, excluding creativity-task-related thoughts. The figure on the right depicts creativity-task-related thoughts, excluding general off-task thoughts. Error bars represent standard errors of the means.

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

**FAKULTÄT FÜR VERHALTENS- UND EMPIRISCHE KULTURWISSENSCHAFTEN**

**Promotionsausschuss der Fakultät für Verhaltens- und Empirische Kulturwissenschaften der Ruprecht-Karls-Universität Heidelberg**
**Doctoral Committee of the Faculty of Behavioural and Cultural Studies of Heidelberg University**

**Erklärung gemäß § 8 (1) c) der Promotionsordnung der Universität Heidelberg für die Fakultät für Verhaltens- und Empirische Kulturwissenschaften**
**Declaration in accordance to § 8 (1) c) of the doctoral degree regulation of Heidelberg University, Faculty of Behavioural and Cultural Studies**

Ich erkläre, dass ich die vorgelegte Dissertation selbstständig angefertigt, nur die angegebenen Hilfsmittel benutzt und die Zitate gekennzeichnet habe.
I declare that I have made the submitted dissertation independently, using only the specified tools and have correctly marked all quotations.

**Erklärung gemäß § 8 (1) d) der Promotionsordnung der Universität Heidelberg für die Fakultät für Verhaltens- und Empirische Kulturwissenschaften**
**Declaration in accordance to § 8 (1) d) of the doctoral degree regulation of Heidelberg University, Faculty of Behavioural and Cultural Studies**

Ich erkläre, dass ich die vorgelegte Dissertation in dieser oder einer anderen Form nicht anderweitig als Prüfungsarbeit verwendet oder einer anderen Fakultät als Dissertation vorgelegt habe.
I declare that I did not use the submitted dissertation in this or any other form as an examination paper until now and that I did not submit it in another faculty.

Vorname Nachname
First name Family name Lena Steindorf

Datum, Unterschrift
Date, Signature ______________________________