EXPLORING THE EFFECTS OF INTERLEAVING ON MIND-WANDERING

A Thesis
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Alyssa Lynn Miller

In Partial Fulfillment of the Requirements
for the Degree of
MASTER OF SCIENCE

Major Department:
Psychology

April 2021

Fargo, North Dakota

North Dakota State University
Graduate School

Title

Exploring the effects of interleaving on mind-wandering

By

Alyssa Lynn Miller

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

MASTER OF SCIENCE

SUPERVISORY COMMITTEE:

Dr. Kathryn Wissman
Chair

Dr. Verlin Hinsz

Dr. Erin Conwell

Dr. Jennifer Momsen

Approved:

July 8th, 2021
Date

Dr. Mark Nawrot
Department Chair

ABSTRACT

Research has shown that interleaving (a study technique where two topics are practiced in an alternating fashion) enhances learning and memory. However, it is unclear whether interleaving impacts the frequency of mind-wandering. The current research explores the extent to which interleaving reduces mind-wandering using a between-subjects design. Participants completed the experiment remotely and learned 40 bird images from 8 distinct families. Bird images were presented in pairs from two different families (e.g., interleaved) or individually from one family (e.g., blocked). After a brief distractor task, participants completed a final classification test. Participants also completed a post-assessment questionnaire regarding the level of difficulty, confidence, and likelihood of missed information throughout the experiment. Results showed no significant differences between the interleaved versus blocked groups for rates of mind-wandering or final test performance. These outcomes have important implications for memory and retention in authentic learning environments.

ACKNOWLEDGMENTS

I would like to acknowledge and thank my committee members, Dr. Erin Conwell, Dr. Jennifer Momsen, Dr. Verlin Hinsz, and my advisor, Dr. Kathryn Wissman for their expertise, support, and guidance. Additionally, I would like to acknowledge and thank Dr. Katherine Duggan for her statistical know-how and encouragement when I needed it. Putting this document together was one of the most challenging and rewarding experiences of my life and I would not have made it through this process without you all by my side. Thank you all so much.

DEDICATION

I dedicate this Master's Thesis to my Mom and Dad, the wonderful educators I've had throughout my educational career (including my advisor – Dr. Kathryn Wissman), and my amazing friends and family who have helped me reach this point in my life. Thank you for your endless support, compassion, late night phone calls, and daily pep talks. I'm truly blessed to be where I am, especially as the first woman in either side of my family to go beyond a Bachelor's degree, and it wouldn't have been possible without all of your help.

TABLE OF CONTENTS

ABSTRACT iii

ACKNOWLEDGMENTS iv

DEDICATION v

LIST OF TABLES viii

LIST OF FIGURES ix

1. INTRODUCTION 1

  1.1. Ways to Reduce Mind-Wandering 2

  1.2. Interleaving as a Learning Technique 4

  1.3. Theoretical Background of Interleaving 5

  1.4. Theoretical Background of Mind-Wandering 9

  1.5. Research Bridge from Interleaving to Mind-Wandering 11

  1.6. Overview of Current Research 13

2. METHOD 15

  2.1. Participants and Design 15

  2.2. Materials 15

  2.3. Procedure 17

  2.4. Scoring 20

3. EXPERIMENT RESULTS 22

  3.1. Rates of Mind-Wandering 22

  3.2. Final Test Performance 23

  3.3. Correlation Between Mind-Wandering and Final-Test Performance 25

  3.4. Post-Assessment Questionnaire Responses 25

4. DISCUSSION 27

REFERENCES 32

APPENDIX. EXAMPLE OF BIRD IMAGES USED IN THE CURRENT EXPERIMENT 42

LIST OF TABLES

Table Page

1. Final Recall Performance by Bird Family 25

2. Post-Assessment Questionnaire Responses Across Groups 26

LIST OF FIGURES

Figure Page

1. Example of Final Classification Test Question....20

2. Proportion of On-Task Thinking Reported for Each of the Five Mind-Wandering Probes (MW1 – 5) and Overall (MW) During Learning....23

3. Final Test Performance for the Interleaved versus Blocked Group....24

1. INTRODUCTION

Students are tasked with learning new information throughout their educational career. Learning the information well requires a substantial amount focus during class and when studying. Unfortunately, students find this to be difficult and report that they mind-wander at a high rate when learning in class. Mind-wandering is a phenomenon described as a shift of attention away from a primary task (Smallwood & Schooler, 2006; Szpunar et al., 2013; Wammes et al., 2016). Given that students tend to mind-wander during learning, their overall understanding and retention of the to-be-learned information can be negatively impacted (Risko et al., 2012; Schacter & Szpunar, 2015; Szpunar et al., 2013b). In an effort to support learning, the current work takes a novel approach of exploring how a well-established learning technique impacts the frequency of mind-wandering. In particular, the current research investigates whether interleaving reduces the frequency of mind-wandering for categorical learning using a between-subjects design.

Mind-wandering is a phenomenon that is a common experience among people, with some research estimating that individuals mind-wander between 30-50% during any given day (Hawkins et al., 2015). As a reminder, mind-wandering refers to a shift of attention away from the task at hand (Smallwood & Schooler, 2006). This shift of attention can happen intentionally or unintentionally, which may depend on a variety of specific individual differences or particular experimental manipulations (Seli et al., 2016a). For example, individuals who have been diagnosed with attention-deficit/hyperactivity disorder (ADHD) or obsessive-compulsive disorder (OCD) are more likely to report higher rates of unintentional mind-wandering (for ADHD, see Bozhilova et al., 2018; McVay & Kane, 2010; Seli et al., 2016a; for OCD, see Seli et al., 2016a). In terms of experimental manipulations, several factors have been shown to induce mind-wandering. For example, having participants engage in restudy (versus retrieval practice) during learning (Peterson & Wissman, 2020; Phillips et al., 2016) and having participants read difficult (versus easy) information (Feng et al., 2013) have both been shown to increase mind-wandering rates. For purposes of the current research, all instances of mind-wandering, whether intentional or unintentional, will be grouped together and categorized as mind-wandering (i.e., off-tasking thinking). This aligns with previous research given that the distinction between unintentional versus intentional mind-wandering remains unclear (Seli et al., 2016b).

1.1. Ways to Reduce Mind-Wandering

Although research has shown a number of different factors that contribute to increased rates of mind-wandering in social and educational settings, research has also explored ways to reduce the likelihood of mind-wandering (for region of proximal learning, see Xu & Metcalfe, 2016; for mindfulness training, see Mrazek et al., 2013; Rahl et al., 2017; for interpolated testing, see Szpunar et al., 2013a). For example, Mrazek et al., (2013) explored whether mindfulness training would improve reading comprehension and the extent to which individuals experience distracting thoughts. Of interest here was the authors exploration of reducing distracting thoughts (i.e., mind-wandering) in students. Participants were randomly assigned to either a mindfulness or nutritional class that met eight times across a two-week time period. During the week prior to the assigned class, participants completed a verbal-reasoning section from the Graduate Record Examination. While completing this task, participants responded to eight probes that were presented at unpredictable and pre-determined (i.e., pseudo-random) intervals in which they reported how on-task or not on-task their attention was using a 5-point response scale (ranging from 1 being “completely on task” to 5 being “completely on unrelated concerns”). Outcomes showed no differences in on-task thinking between the two classes, which indicates the frequency of mind-wandering was similar across all participants prior to the manipulation. Importantly, participants also completed the same verbal-reasoning task and responded to the same mind-wandering probes during the week following completion of the assigned class. Mrazek et al., (2013) found that participants in the mindfulness training class versus the nutritional training class reported being more on-task during the verbal-reasoning section from the Graduate Record Examination in the week following the completion of the class. These findings suggest that mindfulness training is one approach that can be used to reduce mind-wandering rates.

As briefly mentioned above, research has also explored differences in mind-wandering rates when using interpolated testing. For example, Szpunar et al., (2013a) evaluated the extent to which the frequency of mind-wandering is affected when learning during an online lecture; of greatest interest here were mind-wandering rates in the test group versus the restudy group. Participants in Szpunar et al., (2013a; Experiment 2) were tasked with learning an introductory lecture in statistics, which was divided into four segments. During each segment, participants responded to a mind-wandering probe, which asked "Are you mind-wandering? YES or NOCritically, after segments 1-3, participants were either asked to answer questions about key concepts from the lecture (test group) or asked to study questions and provided answers about key concepts from the lecture (restudy group). Results showed that participants reported less mind-wandering in the testing group compared to the restudy group. Indeed, in the testing group, only 19% of responses to the probes indicated mind-wandering (compared to 39% of the probes in the restudy group). Results also showed a negative correlation between low mind-wandering rates and final test performance, suggesting that less mind-wandering facilitates comprehension. Szpunar et al. (2013) findings indicate that incorporating interpolated testing during learning is another factor that can reduce the frequency of mind-wandering. Continuing to explore factors that impact mind-wandering rates -in particular, finding ways to *reduce* the frequency of mind-wandering- has important implications for learning in real-world environments.

1.2. Interleaving as a Learning Technique

Research has explored various learning techniques that facilitate learning and retention for participants, with findings indicating a handful of effective techniques (for a review, see Dunlosky et al., 2013; also see, Weinstein et al., 2018). Of primary interest here is interleaving—a learning technique in which an individual studies two or more topics at the same time in an alternating fashion. To illustrate, imagine a learner was attempting to learning Topic A and Topic B. One approach would be for a student to study the two topics in a blocked fashion (AAABBB). Alternatively, a student could study the two topics in an interleaved fashion (ABABAB). Research has shown that using interleaving (i.e., ABABAB) enhances learning and retention (Birnbaum et al., 2013; Foster et al., 2019; Kang & Pashler, 2012; Kornell & Bjork, 2008; Pan et al., 2019; Wahlheim et al., 2011a; Zulkiply & Burt, 2013), which is important from an educational standpoint. Interleaving benefits have been observed across a variety of delays, learners, and types of to-be-learned material. For example, previous research has established that benefits of interleaving are observed on an immediate test (Birnbaum et al., 2013; Wahlheim et al., 2011a), after one day (Taylor & Rohrer, 2010), after two days (Pan et al., 2019), after three days (Landin et al., 1993), after one week (Foster et al., 2019; Pan et al., 2019), and even after two weeks (Rohrer et al., 2014). Benefits of interleaving have also been found with various types of learners, ranging from young children (Gluckman et al., 2014) to adults over the age of 60 (Bahrick et al., 1993). Regarding the different types of to-be-learned materials, interleaving benefits have been observed for learning musical sequences (Simmons, 2012), foreign languages (Carpenter & Mueller, 2013; Pan et al., 2019), math problems (Foster et al., 2019; Rohrer et al., 2014; Rohrer et al., 2020; Taylor

& Rohrer, 2010), pieces of art (Kang & Pashler, 2012; Kornell & Bjork, 2008; Metcalfe & Xu, 2016; Yan et al., 2016; Yan et al., 2017), and categorical learning (Birnbaum et al., 2013; Hall et al., 1994; Landin et al., 1993; Kornell & Vaughn, 2018; Pani et al., 2013; Wahlheim et al., 2011a). In sum, research has established robust effects for using interleaving as a learning technique, with a recent meta-analysis indicating an effect size of Hedge's *g* = 0.42 (Brunmair & Richter, 2019).

1.3. Theoretical Background of Interleaving

Given the robust nature of the benefits associated with interleaving, researchers have examined underlying mechanisms thought to underlie the effect. The leading mechanism is the *discriminative-contrast hypothesis*, which states that when items from different categories are studied in an alternating format across successive trials, the differences between the two categories are easier to distinguish and in turn enhance performance on a later test (Kang & Pashler, 2012; Kornell & Bjork, 2008). To evaluate the discriminative-contrast hypothesis, Wahlheim et al., (2011a) used a categorical learning task of presenting participants with bird images from different families. Participants were asked to learn 120 images of birds during practice, with 10 exemplars from 12 bird families. The bird images were either presented in a spaced format (i.e., intermixing exemplars from different families) or a blocked format (i.e., separating exemplars by family) The images of the birds were either presented one at a time or in pairs for both the spaced and massed conditions. Therefore, the final design included four groups: spaced-single, spaced-pairs, massed-single, and massed-pairs. On a final classification test, participants were shown both previously studied and novel images of birds and asked to provide the name of the family to which the image of the bird belonged. Wahlheim et al., (2011a) found that final test performance for previously studied and novel exemplars was greater for participants in the spaced-pairs group. , whereas final test performance was lowest for participants in the massed-single group. Findings from Wahlheim et al., (2011a) provide support for the discriminative-contrast hypothesis in that images using interleaving (i.e., side by side presentation of exemplars from different families) facilitated learners' ability to discriminate amongst the different bird families.

Extending on the Wahlheim et al. (2011a) findings, Birnbaum et al., (2013; Experiment 2) further explored the discriminative-contrast hypothesis by investigating the extent to which completing a filler task in between the presentation of exemplars disrupted inductive learning when using interleaving[1]. Participants studied different species of butterflies, with butterfly images being shown one at a time in an interleaved fashion (i.e., alternating between different species). For participants in the contiguous group, participants were shown each butterfly image successively. For participants in the temporally spaced group, participants completed a 10-second filler task in between the presentation of each butterfly image. Following the study of all items, participants completed an inductive-learning test, which asked participants to match new butterfly images with their species. Results showed that participants who studied the butterfly images successively (compared to temporally spaced) performed better on the final inductive-learning test. This pattern of results provides additional evidence for the discriminative-contrast hypothesis and suggests that filler tasks disrupt participants' ability to discriminate between exemplars when utilizing interleaving.

As suggested by Wahlheim et al., (2011a) and Birnbaum et al., (2013), interleaving benefits may derive from the act of discriminating two or more items, which in turn highlights the differences between the to-be-learned items. Discriminating between to-be-learned information may be especially important for math-related content, in which students often need to be able to

---

[1] Experiment 2 from Birnbaum et al. (2013) also included blocked comparison groups, but are not discussed here for purposes of brevity.

both identify and then perform the appropriate strategy for solving a particular math problem. Rohrer et al., (2014) evaluated whether interleaving could facilitate the learning and solving of different kinds of math problems. In the experiment, students were asked to learn four types of math problems; for simplicity, math problems are referred to as A, B, C, and D. For one group of students, problems A and B were interleaved and problems C and D were blocked; for the other group of students, problems A and B were blocked and problems C and D were interleaved. Across a nine-week practice phase, students practiced solving these problems using either interleaving or blocking. Two weeks after their final assignment, students completed a final test in which they were asked to solve novel problems. Outcomes showed final test scores were significantly greater for problems that had been practiced using interleaving versus blocking. These results suggest that interleaving can help learners discriminate among different kinds of problems, which in turn benefits selection and execution of the appropriate strategy. Although prior work evaluating the discriminative contrast hypothesis did not involve mind-wandering, outcomes are indirectly relevant for current purposes. More specifically, if interleaving engenders discrimination and highlights differences between the to-be-learned material, it may help to keep attention. Therefore, it is hypothesized that utilizing interleaving may reduce the frequency of mind-wandering during learning.

Although the discriminative-contrast hypothesis has been the most widely evaluated theoretical mechanism for interleaving benefits, another theoretical account may also explain why interleaving is likely to reduce mind-wandering. The *retrieval-effort hypothesis* states that successful retrieval during practice enhances memory to a greater extent when the retrieval is more versus less effortful (Buckner et al., 1998; Pyc & Rawson, 2009; Schmidtke, 2014; Vaughn et al., 2013). To explore the retrieval-effort hypothesis, Vaughn et al., (2013) examined whether item difficulty played a role in how well a participant can retrieve an item correctly on a final test. In Experiment 2, participants were asked to learn two sets of Lithuanian-English word pairs; 30 easy pairs were shown in one block of practice and 30 difficult pairs were shown in another block of practice. After participants completed the study phase in each respective block, they then progressed to the practice phase. During practice, participants were shown Lithuanian words one at a time and asked to provide the corresponding English word. Participants were asked to recall each word pair correctly a pre-determined number of times before the practice phase was complete. One week later, participants returned to the laboratory to complete a final test. Results showed that target recognition (i.e., participants identifying whether an English word was new or old) was greater for difficult versus easy items. Extrapolating on these findings, it is possible that the retrieval-effort hypothesis may also explain why interleaving helps to reduce mind-wandering. Given that interleaving engenders discrimination amongst the to-be-learned material, using this technique (as compared to blocking) would arguably require greater effort by the learner. If interleaving requires more effort from the learner, then it is likely that interleaving may help keep participants' thoughts on task, which ultimately would result in less mind-wandering and greater performance on a final test.

Although the current experiment is not designed to test any one theoretical mechanism, prior theory-based research provides a basis to help guide predictions in the current research. Prior research has established that interleaving benefits on learning and retention are robust, as effects have been observed across various experimental manipulations including learners, delays, and type of to-be-learned information. In addition to the advantageous effects on memory that interleaving has demonstrated as a learning tool, it is plausible that interleaving also has the added benefit of reducing the frequency of mind-wandering.

1.4. Theoretical Background of Mind-Wandering

Given the frequency at which mind-wandering occurs in our daily lives, research has also examined several theoretical underpinnings of mind-wandering. These theoretical accounts include (but are not limited to), executive resources and executive failure. The *executive-resource hypothesis* states that thoughts related to the experimental task (referred to as task-related) and thoughts unrelated to the experimental task (referred to as task-unrelated) compete for limited executive resources (Feng et al., 2013; Smallwood & Schooler, 2006). Executive resources are components of higher-order processing that enable control of thoughts, behaviors, and actions in regard to achieving a certain goal (Gross & Grossman, 2010). Mind-wandering occurs when task-unrelated thoughts outweigh task-related thoughts.

Xu and Metcalfe (2016) explored the moderating effects between mind-wandering and students' region of proximal learning, which refers to the extent to which the difficulty of a task overlaps with a student's understanding. Participants in Xu and Metcalfe (2016; Experiment 1) were tasked with learning English-Spanish word pairs across three levels of difficulty: easy, medium, and difficult. The Spanish-English word pairs were separated into 12 blocks based on difficulty such that all participants experienced four blocks of each level of difficulty. Important for purposes of mind-wandering, participants were also asked to respond to 12 mind-wandering probes during learning, with four probes for each of the three levels of difficulty. On the mind-wandering probes, participants were asked in indicate whether they were 'MIND WANDERING' or 'ON TASK'. Following learning, all participants completed a final test in which they were shown the English word and asked to recall its Spanish translation. Collapsed across difficulty, results showed that participants reported mind-wandering approximately 36% of the time. Most important for current purposes, results also showed that participants reported significantly less mind-wandering when learning material of medium difficulty, as opposed to more difficult material or more easy material. These results suggest that learning information within an individual's region of proximal learning may help reduce mind-wandering rates. Extending these findings to the executive resource hypothesis, it seems plausible that task-unrelated thoughts will be more likely if the to-be-learned material is incredibly difficult for learners or incredibly easy for learners. And in fact, prior research indicates that mind-wandering rates can be high when learning difficult material (Feng et al., 2013; Xu & Metcalfe, 2016; Experiment 1-2) and easy material (Smallwood & Schooler, 2006; Robertson et al., 1997).

Another theoretical framework examining underlying mechanisms of mind-wandering is the *executive failure hypothesis*, which postulates that mind-wandering happens because the executive control system cannot stop the incessant internal thoughts driving an individual to become unaware and distracted (Pachai et al., 2016; Schooler et al., 2011). The undirected shift in attention (i.e., mind-wandering) results from an unintentional failure of the cognitive system as a whole and a generalized failure at the executive control level (Pachai et al., 2016; McVay & Kane, 2010). Similar to the executive resource hypothesis, in which participants are mind-wandering due to a lack of attentional resources to remain on-task, the executive failure hypothesis also posits that participants may be mind-wandering simply because there are thoughts in their minds that are causing them to mind-wander. For example, if an individual experiences incessant thoughts while studying due to being in a negative mood, the ability to sustain attention on the task at hand may decrease, which would result in executive failure and an increased rate of mind-wandering (Risko et al., 2012; Smallwood et al., 2009). If blocked practice fails to keep the attention of the learner, whereas interleaved practice better maintains attention (by engendering compare-contrast processes) of the learner, then executive failure will be more likely to occur in the blocked group as compared to the interleaved group.

As previously mentioned, the current experiment is not designed to test any one theoretical mechanism of mind-wandering; instead, prior theory-based research provides a basis to help guide predictions in the current research. Based on the aforementioned theoretical mechanisms, I predict that using interleaving will require more attention on part of the learner, which in turn will increase on-task thinking and reduce the frequency of mind-wandering in the current research.

1.5. Research Bridge from Interleaving to Mind-Wandering

Metcalfe and Xu (2016) explored the consequences of mind-wandering as it pertains to students' learning when students studied using blocked versus interleaved practice. Participants were tasked with learning paintings from 24 different artists in a within-subjects design. In their study, 12 artists were assigned to the blocked condition and 12 artists were assigned to the interleaved condition. In the blocked condition, participants studied all paintings from one artist before moving on to the paintings from the next artist. In the interleaved condition, participants studied paintings from each of the 12 artists in an alternating fashion. Critically, participants also responded to 12 mind-wandering probes during learning. After studying the paintings and responding to the mind-wandering probes, participants completed a short distractor task and then took a final inductive learning test. The final test consisted of 48 new paintings (four paintings from each of the twelve artists), and participants were asked to type in the artists' last name for each painting presented. Finally, participants were asked to indicate whether they believed the blocked or interleaved presentation resulted in more mind-wandering and which presentation helped enhance their learning. Results from Metcalfe and Xu (2016) showed that participants reported more mind-wandering in the blocked condition compared to the interleaved condition.

Additionally, results showed that final test performance was greater for the interleaved versus blocked condition. Finally, results showed that a greater number of participants reported feeling as though they mind-wandered more and learned worse when paintings were presented in a massed versus interleaved fashion.

The current research extends the Metcalfe and Xu (2016) by exploring the effects of interleaving on mind-wandering using a between-subjects design. Exploring how interleaving effects mind-wandering using a between-subjects design is important because research has shown that some memory phenomena manifest differently when manipulated between-subjects versus within-subjects (for the generation effect, see McDaniel & Bugg, 2008; Mulligan et al., 2019; Slamecka & Graf, 1978; for the testing effect, see Roediger & Karpicke, 2006; Rowland et al., 2014). For example, the testing effect has been shown to be greater in a between-subjects design but diminish in a within-subjects design (for a meta-analysis, see Rowland et al., 2014). Research in this area has shown that testing oneself over previous information can actually help the learning of subsequent (new) information, which likely helps explain why testing effects are weaker with a within-subject design (Butler, 2010; Dunlosky et al., 2013; Rawson & Dunlosky, 2012; Roediger & Butler, 2011; Roediger & Karpicke, 2006; Roediger & Payne, 1982). Presumably, the same could hold true in the current research. More specifically, if individuals are using interleaving as their sole study technique, then they may be less likely to mind-wander overall compared to if they were switching between interleaving and blocking approaches.

In addition to investigating the how interleaving versus blocking impacts mind-wandering in a between--subjects design, the current research has important implications for practical purposes. In real-world settings, students likely use the same learning technique when studying. In other words, it seems unlikely that a student would choose to study psychology material using an interleaved approach, but biology material using a blocked approach. As such, implementing a between-subjects design in the current research provides a more accurate representation of what students do in their everyday lives. Exploring the effectiveness and reliability of particular study techniques in representative ways is important, as the study techniques that students use can be tied to their success in the classroom (Blumner & Richards, 1997; Credé et al., 2008; Mendezabal, 2013). Investigating whether interleaving has the same effect on mind-wandering when manipulated between- versus within-subject will provide a novel contribution to the literature, with outcomes offering important implications for applied purposes for educators and students alike.

1.6. Overview of Current Research

The goal of the current research is to investigate whether interleaving reduces mind-wandering during learning using a between-subjects design. Given COVID-19 precautions, participants in the current research completed the experiment remotely. Thus, the current research will also provide the opportunity to explore how interleaving and its effect on mind-wandering operate in an authentic learning environment. Participants studied images of birds in an interleaved or blocked fashion. Importantly, during learning all participants were exposed to five mind-wandering probes in which they were asked to report on whether they were on-task or off-task (i.e., mind-wandering). Following a brief distractor task, participants completed a final classification test in which they were shown bird images one at a time and asked to match the image with the correct family. Finally, participants completed a self-report, post-assessment questionnaire which asked them about how difficult it was to stay on task during the experiment, whether they felt they had missed important information due to thinking about other things during the experiment, and how confident they were that they stayed on-task throughout the experiment. I predict that (1) rates of mind-wandering will be lower in the interleaved group versus the blocked group, (2) final test performance will be greater in the interleaved group versus the blocked group, and (3) participants will self-report that it was easier to stay on task in the interleaved versus blocked group.

2. METHOD

2.1. Participants and Design

Participants were 161 undergraduate students recruited from the Psychology Department's participant pool who received course credit upon completion. To align with COVID-19 precautions, participants completed the experiment remotely; participants signed up through the department's SONA site and participated remotely via Qualtrics. Participants were randomly assigned to one of two groups: interleaved versus blocked. The targeted sample size was $n = 122$, based on an a priori power analysis using G*Power 3.1.9.4 (Faul et al., 2009), with power set at .95 and $\alpha = .05$ to detect a medium effect of $d = 0.60$ (based on the effect size observed on mind-wandering outcomes by Metcalfe & Xu, 2016). Data for 36 participants were excluded for having prior bird knowledge ($n = 4$), responding to fewer than four mind-wandering probes ($n = 19$), not completing the entire experiment ($n = 6$), or having their overall time on the experiment exceed 30 minutes ($n = 7$). The final sample included 65 participants in the blocked group and 60 participants in the interleaved group. The average age of the sample was 19 years old ($SE = 0.05$), with 82 females, 42 males, and 1 non-binary persons. This experiment is pre-registered on the Open Science Framework (https://bit.ly/2XgHszB).

2.2. Materials

Materials included images of birds developed by Wahlheim et al., (2011b; also see Birnbaum et al., 2013; Wahlheim et al., 2011a; Wahlheim et al., 2012). Each bird image was shown against a tan background with its species name shown above the image. Bird images belonged to one of eight families: Orioles, Warblers, Finches, Swallows, Vireos, Grosbeaks, Thrushes, and Chickadees. An example bird image from each family is provided in the Appendix. The bird families were presented in the following order: Orioles, Warblers, Finches, Swallows, Vireos, Grosbeaks, Thrushes, and Chickadees. The bird families and exemplars from within each family were chosen based on item difficulty from prior research using this material set. More specifically, the bird images were chosen to equate item difficulty across the bird families, with previous research showing correct classification at approximately 55% (Wahlheim et al., 2011b). Five bird images from each family were presented for a total of 40 bird images. In addition, 10 bird images were used as buffer items at the start of the experiment, which allowed participants to get familiarized with the experiment before going to the to-be-learned information and to circumvent primacy effects; these buffer bird images were selected from two, different bird families (Flycatchers and Sparrows).

The materials also included a mind-wandering probe (adopted from Unsworth & Robinson, 2016; also see, Peterson & Wissman, 2020). The mind-wandering probe asked, "What were you just thinking about?" and participants were asked to choose from the following options: **(1)** task, **(2)** task performance, **(3)** everyday stuff, **(4)** current state of being, **(5)** personal worries, **(6)** daydreams, and **(7)** other.

Finally, materials included a self-report questionnaire which probed participants about their ability to stay on task during the experiment. More specifically, participants were asked the following three questions: **(1)** How difficult was it for you to stay on task during the experiment; participants selected from five response options ranging from the lowest extreme (e.g., "Very easy") to the highest extreme (e.g., "Very difficult"); **(2)** During today's experiment I missed important pieces of information because I was thinking of other things; participants selected from five response options ranging from the lowest extreme (e.g., "Strongly disagree") to the highest extreme (e.g., "Strongly agree"); and **(3)** How confident do you feel that you stayed on task for today's experiment; participants selected from five response options ranging from the lowest extreme (e.g., "Extremely unconfident") to the highest extreme (e.g., "Extremely confident").

2.3. Procedure

The current study was completed remotely. Before beginning the experiment, participants responded to a screening question to help ensure that the participants did not have pre-existing or extensive knowledge about birds and ornithology. Participants were asked, "Do you have extensive knowledge about birds and ornithology? (For example, can you easily identify birds?)," to which they were prompted to respond either "yes" or "no." If a participant selected "yes", they were directed to the end of the experiment and told they were not eligible for the current experiment due to their bird knowledge. If a participant selected "no" to the Bird Knowledge Question, they were directed to the informed consent form. Participants indicated consent to participate in the experiment by selecting "I agree" or "I decline". All participants then read the same instructions, which gave an overview of what they were expected to do in the experiment. The instructions were as follows:

> "During today's experiment, you will be asked to learn species from several different bird families. You will be shown pictures of birds for a set amount of time and the computer will automatically advance you when time is up. You should do your best to try and remember all of the birds. Please do not take notes or write anything down during today's experiment.
>
> In this experiment, we are also interested in studying mind-wandering. Mind-wandering refers to when an individual is not focusing on the task at hand (in this case, learning the birds) and thinking about something else. Throughout the experiment you will also randomly be asked to report if you are mind-wandering. Given that we are interested in studying mind-wandering, please be honest when you are asked to report whether you are engaging in mind-wandering or focused on the task.
>
> For the mind-wandering probes, you will be asked "What were you just thinking about?" and given seven options from which to choose:

Task - thinking about the stimulus words or an appropriate response (for example, you were focused on studying the birds)
Task performance - evaluating your performance (for example, you were thinking about how well you were learning the birds)
Everyday stuff - thinking about recent or impending life events or tasks (for example, you were thinking about errands you need to run)
Current state of being - thinking about conditions (for example, you were thinking about feeling hungry)
Personal worries - thinking about concerns, troubles, or fears (for example, you were thinking about how you need to pay your bills this month)
Daydreams - having fantasies disconnected from reality (for example, you were thinking about taking a trip to the beach)
Other - other thought types

For each probe, you will be given 10 seconds to provide a response. Throughout the experiment, everything (including presentation of birds) will be paced by the computer. The computer will automatically advance you when time is up. As a reminder, mind-wandering is a normal phenomenon and please be honest when you are asked to report whether you are mind wandering or on task.

When you are ready, please begin the experiment."

To learn the bird families, participants studied bird images in the pre-determined order (discussed under Materials). Participants were asked to study bird images according to one of two study schedules: interleaved or blocked. In the interleaved group, participants were presented with two bird images belonging to two different bird families. For example, participants were shown one image from the Oriole family and one image from the Warbler family to study. Each pairing was presented for 16 seconds. After 16 seconds had elapsed, participants then studied the next image from the Oriole family and the next image from the Warbler family. After the five images from the first two bird families (i.e., Orioles and Warblers) had been shown, images from the next two bird families (i.e., Finches and Swallows) were shown to participants in the same manner. Presentation of all bird images and families followed this format for the interleaved group. In the blocked group, participants were presented with one bird image from one family. For example, participants were shown one image from the Oriole family to study. The image was presented for

8 seconds. After 8 seconds had elapsed, participants then studied the next image from the Oriole family. After the five images from the first bird family (i.e., Orioles) had been shown, images from the next bird family (i.e., Warblers) were shown to participants in the same manner. Presentation of all bird images and families followed this format for the blocked group.

While studying the bird images, participants were also asked to respond to five mind-wandering probes. To ensure that both the interleaved group and the blocked group provided mind-wandering ratings at the same time, each mind-wandering probe occurred at a pre-determined time. However, the times of the probes were not regularly spaced so that participants could not anticipate them (Robison et al., 2019). Mind-wandering probes occurred at 112 seconds, 186 seconds, 242 seconds, 348 seconds, and 390 seconds for both the interleaved and blocked group. Participants were given 10 seconds to respond to each mind-wandering probe.

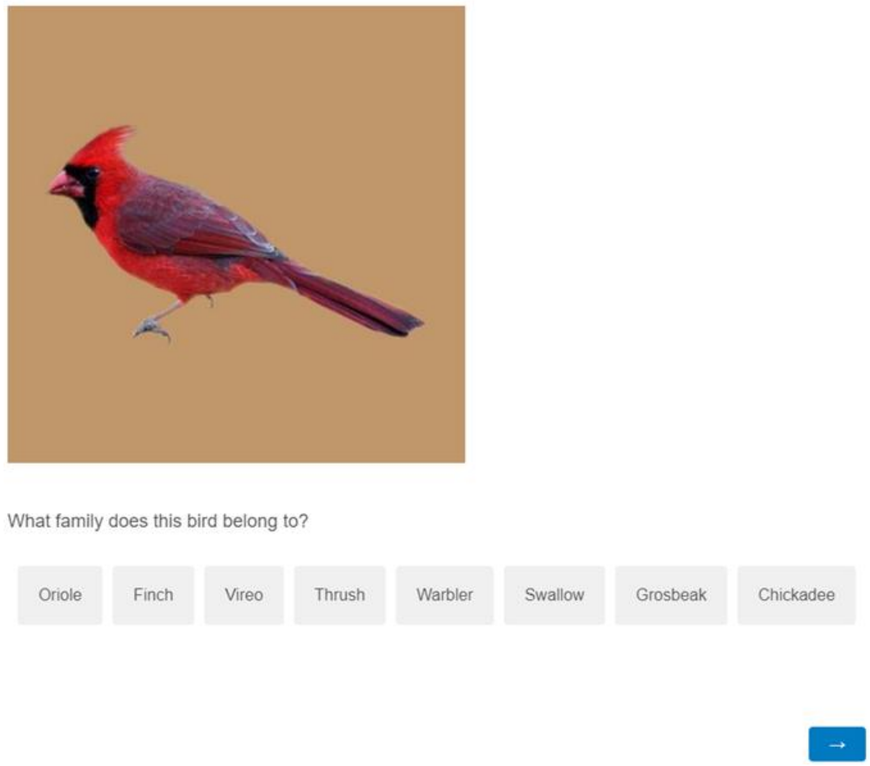
After all of the bird images were presented, participants were then asked to complete an unrelated distractor task for three minutes. During this task, participants were asked to create words using the letters in the words NEVERTHELESS and UPPERCLASSMAN; for each word, participants were allotted 90 seconds to generate as many words as possible. Once three minutes had elapsed, participants completed the final test. Instructions for the final test were as follows:

> "You will now take a final classification test over the bird families you studied today. On this test, you will be shown bird pictures one at a time and asked to indicate which family the bird belongs to. Once you have selected your response for each bird picture, please click the arrow to advance yourself to the next bird picture."

On the final test, participants were shown each previously-studied bird image one at a time in random order. Below each image, the names of the eight primary families were listed and participants were asked to match the bird images to its family (see Figure 1).

Figure 1

Example of Final Classification Test Question



Note. In this example, the correct answer is Grosbeak.

Participants were given as much time as they needed to match the bird image to the family and feedback was not provided. After completing the classification test, participants answered three questions in the post-assessment questionnaire (described under Materials). Once the participants responded to the post-assessment questionnaire, they were asked several demographic questions; participants were not required to disclose demographic information. Finally, participants read a debriefing statement and thanked for their participation.

2.4. Scoring

For the mind-wandering probe, responses "task" or "task performance" were coded as 1 (on-task thinking) and all other responses were coded as 0 (mind-wandering). Mind-wandering probe responses were summed and are reported as a proportion score, with higher scores indicating more on-task thinking. The classification test questions were scored as 0 (incorrect classification) or 1 (correct classification), for a total of 40 points possible. Overall scores were averaged and are reported as proportion correct. Finally, post-questionnaire responses were coded with values ranging from 1 (lowest extreme) to 5 (highest extreme), with higher scores indicating responses consistent with greater mind-wandering. Post-assessment questionnaire responses are reported in two ways. First, responses for each of the three questions are averaged across the interleaved and blocked groups and reported as overall means. Second, responses for the three questions are summed and averaged to create a single composite score for each participant; results report the mean composite score for the interleaved and the blocked groups.

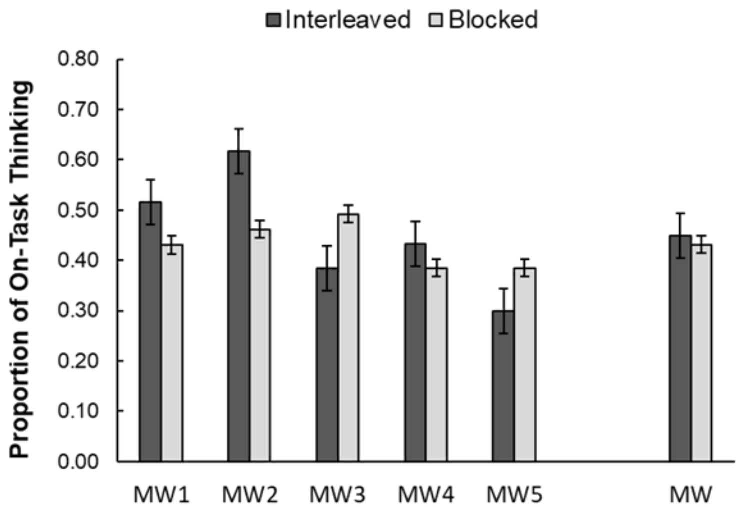
3. EXPERIMENT RESULTS

3.1. Rates of Mind-Wandering

In contrast to predictions, overall rates of mind-wandering did not differ for participants in the interleaved group versus the blocked group. As shown in the right side of Figure 2, rates of on-task thinking were similar in the interleaved group (M = 0.45, SE = 0.03) versus the blocked group (M = 0.43, SE = 0.03), $t(123) = -0.43$, $p = 0.668$. To investigate outcomes further, on-task thinking for each of the five mind-wandering probes was examined (see the left side of Figure 2). Interestingly, the interleaved versus blocked group showed numerically more on-task thinking early in the experiment, whereas this difference was less pronounced (and even reversed) later in the experiment. Independent t-tests for each of the five mind-wandering probes did not cross the threshold for statistical significance, but it is interesting to note the pattern of results changed throughout the course of the experiment. As such, an exploratory follow-up 2 (Group: interleaved vs. blocked) × 5 (Mind-wandering Probe: 1 vs. 2 vs. 3 vs. 4 vs. 5) repeated measures ANOVA was conducted. Outcomes revealed a main effect of mind-wandering, such that mind-wandering rates were higher for later probes as compared to earlier probes, $F(4,120) = 3.11$, $p = 0.018$. This suggests that overall mind-wandering increased across the experiment, regardless of group (for similar outcomes, see Metcalfe & Xu, 2016). The main effect of group and the interaction between group and mind-wandering was not significant, (Fs > 0.185 and ps > 0.136).

Figure 2

Proportion of On-Task Thinking Reported for Each of the Five Mind-Wandering Probes (MW1 – 5) and Overall (MW) During Learning



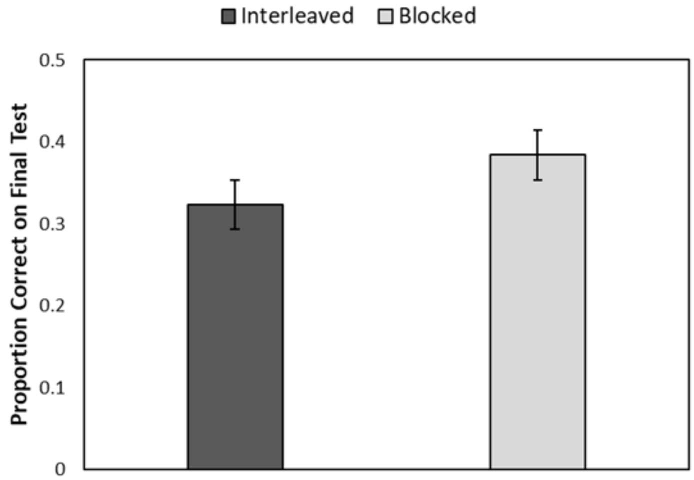
Note. Error bars represent the standard error of the mean.

3.2. Final Test Performance

Overall performance on the final classification test was quite low—approximately 36%; this outcome is discussed further in the Discussion. Interestingly, participants showed no difference on the final classification test (see Figure 3), such that performance was similar in the blocked group (*M* = 0.39, *SE* = .03) and the interleaved group (*M* = 0.32, *SE* = 0.02), *t*(123) = 1.73, *p* = 0.086.

Figure 3

Final Test Performance for the Interleaved versus Blocked Group



Note. Error bars represent the standard error of the mean.

In fact, results indicate a numerical trend in the *opposite* direction of an interleaving benefit; this surprising outcome is discussed further in the Discussion. Based on results from a mixed effects logistic regression, performance for seven out of the eight bird families was numerically lower for participants in the interleaved group than participants in the blocked group (β = -0.31, *SE* = 0.19, *Z*(5000) = -1.70, *p* = 0.090). The only family on the final classification test that the interleaved group had higher performance (*M* = 0.38, *SE* = 0.04) than the blocked group (*M* = 0.25, *SE* = 0.04) on was the Warbler family, *t*(123) = -2.26, *p* = 0.026. Final classification performance is broken down by each of the eight families and provided in Table 1 for interested readers.

Table 1

Final Recall Performance by Bird Family

	Oriole	Warbler	Finches	Swallows	Vireos	Grosbeaks	Thrushes	Chickadees	Overall
Interleaved	0.50 (0.04)	0.38 (0.04)	0.29 (0.04)	0.31 (0.04)	0.17 (0.03)	0.30 (0.03)	0.27 (0.04)	0.38 (0.04)	0.32 (0.02)
Blocked	0.66 (0.03)	0.25 (0.04)	0.35 (0.04)	0.42 (0.04)	0.21 (0.03)	0.36 (0.04)	0.35 (0.04)	0.48 (0.04)	0.38 (0.03)

Note. Final recall performance broken down by bird family during final test. The values in parentheses represent standard error (SE) values.

3.3. Correlation Between Mind-Wandering and Final-Test Performance

Across all of the participants, regardless of group assignment, there was a positive correlation between participants' on-tasking thinking and performance on the final classification test, $r = 0.308$, $p < 0.001$. More specifically, more on-task thinking was associated with higher final test performance. The correlation between proportion of on-task thinking and final test performance was significant for participants in the interleaved group ($r = .261$, $p = .044$) and participants in the blocked group ($r = .362$, $p = .003$).

3.4. Post-Assessment Questionnaire Responses

Table 2 provides outcomes from the post-assessment questionnaire. Participants in the interleaved versus blocked group showed no differences in reported rates of difficulty staying on-task, confidence in staying on task, or feelings of missed information (all *ps* > .450). All three questions on the post-assessment questionnaire were significantly correlated with one another (all *ps* < 0.01). As such, to enhance the reliability and validity of the post-assessment measure, a composite score was computed. Participants in the interleaved versus blocked group showed no differences in their composite scores, $t(123) = -0.181$, $p = 0.856$. Although the post-assessment questionnaire outcomes may not be particularly exciting, they do provide some indication of metacognitive accuracy; indeed, mind-wandering rates during learning were not different between the two groups and responses to post-assessment questions map on to that finding.

Table 2

Post-Assessment Questionnaire Responses Across Groups

	Difficulty	Missed Information	Confidence	Composite Score
Interleaved	3.77 (0.10)	3.93 (0.12)	3.32 (0.12)	3.65 (0.11)
Blocked	3.65 (0.13)	3.85 (0.12)	3.45 (0.12)	3.67 (0.09)

Note. During the post-assessment questionnaire, the rate of extremity for each question is reported. The values in parentheses represent standard error (SE) values.

4. DISCUSSION

The current research was designed to explore whether interleaving decreased rates of mind-wandering when learning categorical information in a between-subjects design. Three noteworthy patterns emerged from the current research. First, there was no statistical difference in rates of mind-wandering for the interleaved versus blocked groups during learning. Despite evidence of a slight numerical advantage for more on-task thinking in the interleaved group early in learning, these outcomes did not reach statistical significance and ultimately this advantage disappeared. Second, there was no statistical difference on final classification test performance for the interleaved versus blocked groups. Given that prior research has shown interleaving to often have robust effects on learning and retention, this outcome was particularly surprising (and is discussed in more detail below). Finally, no differences emerged on the post-assessment questionnaire between the interleaved and blocked groups. Although unexpected, results from the current research suggest that interleaving may not reduce the frequency of mind-wandering and have important implications for learning in real-world environments. Below, I discuss possible reasons as to why such results emerged and possible ideas for future research.

As a critical first step, future research should attempt to replicate the observed findings. There has been a recent emphasis on the importance of replicability in the field of psychology, particularly for novel effects (Camerer et al., 2018; Colling & Szűcs, 2021; Fabrigar et al., 2020; Field et al., 2019; McElreath & Smaldino, 2015; Zwaan et al., 2018). Indeed, replication is important to establish the reliability and generalizability of effects. Given that the current research is the first exploration of how interleaving impacts the frequency of mind-wandering using a between-subjects design, a direct replication is warranted. More specifically, a follow-up experiment that includes the same experimental design and methodology would provide additional data, which would strengthen conclusions that can be made about how interleaving impacts mind-wandering using a between-subjects design and in an authentic learning environment. In addition to attempting to replicate the observed effects, future research should also aim to extend on the current research. Several potential extensions on the current work are discussed throughout the remainder of the Discussion.

Why did interleaving not lead to a reduction in mind-wandering during learning? One reason may be that interleaving impacts mind-wandering differently when used in a between-subjects versus within-subjects design. To revisit, Metcalfe and Xu (2016) observed decreased rates of mind-wandering when participants studied paintings in an interleaved versus blocked format. Upon closer examination of these outcomes, Metcalfe and Xu (2016) found that mind-wandering rates were similar for the interleaved versus blocked conditions across the first three blocks of practice; it was only in the fourth (and final) block of practice that interleaving resulted in a reduction in mind-wandering. Their results suggest that using interleaving (as compared to blocking) helped maintain learners' attention over a longer amount of time. Of course, the opposite was observed in the current research, such that the interleaved group showed numerically more on-task at the beginning of the experiment (Probes 1 and 2) but became increasingly off-task as the experiment progressed (Probes 3-5). One possible explanation for this different pattern of results is that when interleaving is manipulated within-subjects the technique helps maintain attention over time, whereas when interleaving in manipulated between-subjects the technique helps capture attention at the outset but not sustain attention. Future research examining the extent to which interleaving may differentially capture and keep attention in within- versus between-subjects designs would be an enlightening direction for future research.

A second possible reason why interleaving may not have led to less mind-wandering concerns how the to-be-learned material was presented in the current experiment. Prior research has evaluated the effects of interleaving using both simultaneous presentation (i.e., presenting two images at the same time from different categories) and successive presentation (i.e., presenting one image at a time alternating between categories), with outcomes showing mixed results for which presentation style is most effective (for simultaneous presentation benefits, see Birnbaum et al., 2012; Kang & Pashler, 2012; Experiment 2; Wahlheim et al., 2011; for successive presentation benefits, see Kang & Pashler, 2012; Experiment 1; Kornell & Bjork, 2008). In terms of interleaving and its relation to mind-wandering, perhaps presenting two images at the same time is too much information for participants to process. More specifically, being shown two bird images simultaneously may overwhelm executive resources and in turn increase the likelihood of mind-wandering. To explore this possibility, future research could manipulate successive versus simultaneous presentation of the bird images and examine mind-wandering rates. For example, in an interleaving-simultaneous group, two bird images from two different families could be presented at the same time (identical to the interleaving group in the current research). In contrast, in an interleaving-successive group, bird images could be shown one at a time but alternate between two different families (e.g., the first image of an Oriole followed by the first image of a Warbler, then the second image of an Oriole followed by the second image of a Warbler, and so on). If presenting one image at a time is less likely to overwhelm executive resources and engender greater focus on the task at hand (Feng et al., 2013; Smallwood & Schooler, 2006), then rates of mind-wandering would be reduced in the interleaving-successive group versus the interleaved-simultaneous. Future research examining how simultaneous versus successive presentation during interleaving impacts the frequency of mind-wandering would be informative.

A third reason why interleaving may not have led to less mind-wandering is because learners completed the experiment outside of a controlled environment. Participating via remote environments may contribute to high levels of fatigue and cause participants to fall into a state of mindlessness, as opposed to if they were in a more controlled setting (Labrague & Ballad, 2020; Thomson et al., 2015). Furthermore, given that a majority of students have been learning remotely and/or in a hybrid format, it is possible that high levels of fatigue may have been prevalent at the outset of this experiment. Indeed, when individuals are in front of a screen for prolonged periods of time, thoughts about technology begin to drive mind-wandering and lead to a loss of attention (Hollis & Was, 2016). Furthermore, participants may have been in an environment in which common distractions in their surroundings may have impacted rates of mind-wandering (Dost et al., 2020). Interestingly, a majority of prior research investigating mind-wandering has occurred in much more controlled settings (for exceptions, see Hollis et al., 2016; Schacter & Szpunar, 2015; Szpunar et al., 2013b). Given the recent shift to remote learning, it seems plausible that the frequency of mind-wandering may operate differently in these types of settings. Further examination of how mind-wandering rates differ in more controlled versus real-world environments is a fruitful avenue for future research.

In addition to having no effect on mind-wandering rates, the current outcomes revealed that interleaving produced no memory benefit on the final classification test. One possible reason why an interleaving benefit was not observed on the final test concerns the difficultly of the to-be-learned material. Although the material set was intentionally selected to be of medium-level difficulty (based on norming data from Wahlheim et al., 2011), it is clear that participants struggled to learn this material. Indeed, collapsed across the two groups, correct classification on the final test was only 36%. These outcomes suggest that additional study of the to-be-learned information was needed. In other words, simply studying the bird images using an interleaved fashion was not sufficient for learning and retention. Perhaps combining interleaving with another learning technique (e.g., retrieval practice) would yield a memory benefit. Beyond difficulty of learning in particular, the low final test performance also offers another reason as to why a reduction in mind-wandering was not observed—namely, the to-be-learned material was outside of participants' region of proximal learning (see Xu & Metcalfe, 2016). Indeed, prior research suggests that mind-wandering rates can increase when material is difficult, which was clearly the case here.

At a more general level, the current outcomes have direct implications for the ecological validity of using interleaving. The current research took place in a remote setting (i.e., not in the laboratory); given the surprising findings, additional exploration of this technique and how it impacts the frequency of mind-wandering and memory in authentic learning environments seems necessary. More specifically, future research exploring the consequences of interleaving in real-world environments will inform recommendations given to instructors and students.

The current research provides a novel exploration of how interleaving affects mind-wandering using a between-subjects design when learning categorical information. Outcomes provide preliminary evidence that interleaving may not reduce the frequency of mind-wandering when manipulated between-subjects and used in a more representative learning environment, but more research is needed. Further examination of interleaving and its effects on mind-wandering will be important for practical recommendations given to educators and students.

REFERENCES

Bahrick, H. P., Bahrick, L. E., Bahrick, A. S., & Bahrick, P. E. (1993). Maintenance of foreign language vocabulary and the spacing effect. *Psychological Science, 4*(5), 316-321.

Birnbaum, M. S., Kornell, N., Bjork, E. L., & Bjork, R. A. (2013). Why interleaving enhances inductive learning: The roles of discrimination and retrieval. *Memory & Cognition*, *41*(3), 392-402.

Blumner, H. N., & Richards, H. C. (1997). Study habits and academic achievement of engineering students. *Journal of Engineering Education*, *86*(2), 125-132.

Brunmair, M., & Richter, T. (2019). Similarity matters: A meta-analysis of interleaved learning and its moderators. *Psychological Bulletin, 145*(11), 1029-1052.

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

Buckner, R. L., Koutstaal, W., Schacter, D. L., Wagner, A. D., & Rosen, B. R. (1998). Functional–anatomic study of episodic retrieval using fMRI: I. Retrieval effort versus retrieval success. *Neuroimage*, *7*(3), 151-162.

Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(5), 1118-1133.

Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T. H., Huber, J., Johannesson, M., ... & Wu, H. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behaviour*, *2*(9), 637-644.

Carpenter, S. K., & Mueller, F. E. (2013). The effects of interleaving versus blocking on foreign language pronunciation learning. *Memory & Cognition*, *41*(5), 671-682.

Carvalho, P. F., & Goldstone, R. L. (2014). Putting category learning in order: Category structure and temporal arrangement affect the benefit of interleaved over blocked study. *Memory & Cognition*, *42*(3), 481-495.

Carvalho, P. F., & Goldstone, R. L. (2017). The sequence of study changes what information is attended to, encoded, and remembered during category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(11), 1699-1719.

Chen, O., Paas, F., & Sweller, J. (2021). Spacing and Interleaving Effects Require Distinct Theoretical Bases: a Systematic Review Testing the Cognitive Load and Discriminative-Contrast Hypotheses. *Educational Psychology Review*, 1-24.

Credé, M., & Kuncel, N. R. (2008). Study habits, skills, and attitudes: The third pillar supporting collegiate academic performance. *Perspectives on Psychological Science*, *3*(6), 425-453.

Dost, S., Hossain, A., Shehab, M., Abdelwahed, A., & Al-Nusair, L. (2020). Perceptions of medical students towards online teaching during the COVID-19 pandemic: a national cross-sectional survey of 2721 UK medical students. *BMJ Open*, *10*(11), e042378.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, *14*(1), 4-58.

Fabrigar, L. R., Wegener, D. T., & Petty, R. E. (2020). A validity-based framework for understanding replication in psychology. *Personality and Social Psychology Review*, *24*(4), 316-344.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149-1160.

Feng, S., D'Mello, S., & Graesser, A. C. (2013). Mind wandering while reading easy and difficult texts. *Psychonomic Bulletin & Review*, *20*(3), 586-592.

Field, S. M., Hoekstra, R., Bringmann, L., van Ravenzwaaij, D., & Savalei, V. (2019). When and why to replicate: As easy as 1, 2, 3?. *Collabra: Psychology, 5*(1), 1-15.

Firth, J., Rivers, I., & Boyle, J. (2021). Context and Implications Document for: A systematic review of interleaving as a concept learning strategy. *Review of Education*, *9*(2), 685–688. https://doi-org.ezproxy.lib.ndsu.nodak.edu/10.1002/rev3.3265

Foster, N. L., Mueller, M. L., Was, C., Rawson, K. A., & Dunlosky, J. (2019). Why does interleaving improve math learning? The contributions of discriminative contrast and distributed practice. *Memory & Cognition*, *47*(6), 1088-1101.

Gluckman, M., Vlach, H. A., & Sandhofer, C. M. (2014). Spacing simultaneously promotes multiple forms of learning in children's science curriculum. *Applied Cognitive Psychology, 28*(2), 266-273.

Gross, R. G., & Grossman, M. (2010). Executive resources. *Continuum (Minneapolis, Minn.)*, *16*(4 Behavioral Neurology), 140–152. https://doi.org/10.1212/01.CON.0000368266.46038.0e

Hall, K. G., Domingues, D. A., & Cavazos, R. (1994). Contextual interference effects with skilled baseball players. *Perceptual and Motor Skills, 78*(3), 835-841.

Hawkins, G. E., Mittner, M., Boekel, W., Heathcote, A., & Forstmann, B. U. (2015). Toward a model-based cognitive neuroscience of mind wandering. *Neuroscience, 310,* 290-305.

Heijltjes, A., Van Gog, T., & Paas, F. (2014). Improving students' critical thinking: Empirical support for explicit instructions combined with practice. *Applied Cognitive Psychology*, *28*(4), 518-530.

Hollis, R. B., & Was, C. A. (2016). Mind wandering, control failures, and social media distractions in online learning. *Learning and Instruction*, *42*, 104-112.

Kang, S. H., & Pashler, H. (2012). Learning painting styles: Spacing is advantageous when it promotes discriminative contrast. *Applied Cognitive Psychology*, *26*(1), 97-103.

Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories: Is spacing the "enemy of induction"?. *Psychological Science*, *19*(6), 585-592.

Kornell, N. (2009). Optimising learning using flashcards: Spacing is more effective than cramming. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *23*(9), 1297-1317.

Kornell, N., & Vaughn, K. E. (2018). In inductive category learning, people simultaneously block and space their studying using a strategy of being thorough and fair. *Archives of Scientific Psychology*, *6*(1), 138-147.

Landin, D. K., Hebert, E. P., & Fairweather, M. (1993). The effects of variable practice on the performance of a basketball skill. *Research Quarterly for Exercise and Sport*, *64*(2), 232-237.

McDaniel, M. A., & Bugg, J. M. (2008). Instability in memory phenomena: A common puzzle and a unifying explanation. *Psychonomic Bulletin & Review*, *15*(2), 237-255.

McElreath, R., & Smaldino, P. E. (2015). Replication, communication, and the population dynamics of scientific discovery. *PloS One*, *10*(8), e0136088.

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

Mendezabal, M. J. N. (2013). Study habits and attitudes: The road to academic success. *Open Science Repository Education*, (open-access), e70081928.

Metcalfe, J., & Xu, J. (2016). People mind wander more during massed than spaced inductive learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(6), 978-984.

Mrazek, M. D., Franklin, M. S., Phillips, D. T., Baird, B., & Schooler, J. W. (2013). Mindfulness training improves working memory capacity and GRE performance while reducing mind wandering. *Psychological Science*, *24*(5), 776-781.

Mulligan, N. W., Smith, S. A., & Buchin, Z. L. (2019). The generation effect and experimental design. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(8), 1422-1431.

Nakata, T., & Suzuki, Y. (2019). Mixing grammar exercises facilitates long-term retention: Effects of blocking, interleaving, and increasing practice. *The Modern Language Journal*, *103*(3), 629-647.

Pachai, A. A., Acai, A., LoGiudice, A. B., & Kim, J. A. (2016). The mind that wanders: Challenges and potential benefits of mind wandering in education. *Scholarship of Teaching and Learning in Psychology, 2*(2), 134–146. https://doi.org/10.1037/stl0000060

Pan, S. C., Tajran, J., Lovelett, J., Osuna, J., & Rickard, T. C. (2019). Does interleaved practice enhance foreign language learning? The effects of training schedule on Spanish verb conjugation skills. *Journal of Educational Psychology*, *111*(7), 1172-1188.

Peterson, D. J., & Wissman, K. (2020). Using tests to reduce mind-wandering during learning review. *Memory*, *28*(4), 582-587.

Phillips, N. E., Mills, C., D'Mello, S., & Risko, E. F. (2016). On the influence of re-reading on mind wandering. *Quarterly Journal of Experimental Psychology*, *69*(12), 2338-2357.

Pyc, M. A., & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory?. *Journal of Memory and Language*, *60*(4), 437-447.

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 (Washington, D.C.)*, *17*(2), 224–230.

Randall, J. G., Oswald, F. L., & Beier, M. E. (2014). Mind-wandering, cognition, and performance: a theory-driven meta-analysis of attention regulation. *Psychological Bulletin*, *140*(6), 1411-1431.

Rawson, K. A., & Dunlosky, J. (2012). When is practice testing most effective for improving the durability and efficiency of student learning?. *Educational Psychology Review*, *24*(3), 419-435.

Richland, L. E., Bjork, R. A., Finley, J. R., & Linn, M. C. (2005). Linking cognitive science to education: Generation and interleaving effects. In *Proceedings of the Twenty-Seventh Annual Conference of the Cognitive Science Society* (Vol. 27, pp. 1850-55).

Risko, E. F., Anderson, N., Sarwal, A., Engelhardt, M., & Kingstone, A. (2012). Everyday attention: Variation in mind wandering and memory in a lecture. *Applied Cognitive Psychology*, *26*(2), 234-242.

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*(6), 747-758.

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.

Roediger III, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, *15*(1), 20-27.

Roediger III, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*(3), 249-255.

Roediger, H. L., & Payne, D. G. (1982). Hypermnesia: The role of repeated testing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *8*(1), 66.

Rohrer, D., Dedrick, R. F., & Burgess, K. (2014). The benefit of interleaved mathematics practice is not limited to superficially similar kinds of problems. *Psychonomic Bulletin & Review*, *21*(5), 1323-1330.

Rohrer, D., Dedrick, R. F., Hartwig, M. K., & Cheung, C. N. (2020). A randomized controlled trial of interleaved mathematics practice. *Journal of Educational Psychology*, *112*(1), 40-52.

Rowland, C. A., Littrell-Baez, M. K., Sensenig, A. E., & DeLosh, E. L. (2014). Testing effects in mixed-versus pure-list designs. *Memory & Cognition*, *42*(6), 912-921.

Schacter, D. L., & Szpunar, K. K. (2015). Enhancing attention and memory during video-recorded lectures. *Scholarship of Teaching and Learning in Psychology*, *1*(1), 60-71.

Schmidtke, J. (2014). Second language experience modulates word retrieval effort in bilinguals: evidence from pupillometry. *Frontiers in Psychology*, *5*, 137.

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.

Seli, P., Risko, E. F., Smilek, D., & Schacter, D. L. (2016a). Mind-wandering with and without intention. *Trends in Cognitive Sciences*, *20*(8), 605-617.

Seli, P., Risko, E. F., & Smilek, D. (2016b). On the necessity of distinguishing between unintentional and intentional mind wandering. *Psychological Science*, *27*(5), 685-691.

Simmons, A. L. (2012). Distributed practice and procedural memory consolidation in musicians' skill learning. *Journal of Research in Music Education*, *59*(4), 357-368.

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory*, *4*(6), 592-604.

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

Szpunar, K. K., Khan, N. Y., & Schacter, D. L. (2013a). Interpolated memory tests reduce mind wandering and improve learning of online lectures. *Proceedings of the National Academy of Sciences*, *110*(16), 6313-6317.

Szpunar, K. K., Moulton, S. T., & Schacter, D. L. (2013b). Mind wandering and education: from the classroom to online learning. *Frontiers in Psychology*, *4*, Article 495.

Taylor, K., & Rohrer, D. (2010). The effects of interleaved practice. *Applied Cognitive Psychology*, *24*(6), 837-848.

Thomson, D. R., Besner, D., & Smilek, D. (2015). A resource-control account of sustained attention: Evidence from mind-wandering and vigilance paradigms. *Perspectives on Psychological Science, 10*(1), 82-96.

Unsworth, N., & Robison, M. K. (2016). Pupillary correlates of lapses of sustained attention. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(4), 601–615

Vaughn, K. E., Rawson, K. A., & Pyc, M. A. (2013). Repeated retrieval practice and item difficulty: Does criterion learning eliminate item difficulty effects?. *Psychonomic Bulletin & Review*, *20*(6), 1239-1245.

Wahlheim, C. N., Dunlosky, J., & Jacoby, L. L. (2011a). Spacing enhances the learning of natural concepts: An investigation of mechanisms, metacognition, and aging. *Memory & Cognition*, *39*(5), 750-763.

Wahlheim, C. N., Finn, B., & Jacoby, L. L. (2012). Metacognitive judgments of repetition and variability effects in natural concept learning: Evidence for variability neglect. *Memory & Cognition*, *40*(5), 703-716.

Wahlheim, C. N., Teune, R. K., & Jacoby, L. L. (2011b). Birds as natural concepts: A set of pictures from the Passeriformes order. *Retrieved from http://psych. wustl. edu/amcclab/AMCC% 20Materials. htm*.

Wammes, J. D., Seli, P., Cheyne, J. A., Boucher, P. O., & Smilek, D. (2016). Mind wandering during lectures II: Relation to academic performance. *Scholarship of Teaching and Learning in Psychology*, *2*(1), 33-48.

Was, C. A., Hollis, R. B., & Dunlosky, J. (2019). Do students understand the detrimental effects of mind wandering during online learning?. *Computers & Education*, *135*, 113-122.

Weinstein, Y., Madan, C. R., & Sumeracki, M. A. (2018). Teaching the science of learning. *Cognitive Research: Principles and Implications*, *3*(1), Article 2.

Wissman, K. T., Rawson, K. A., & Pyc, M. A. (2012). How and when do students use flashcards?. *Memory*, *20*(6), 568-579.

Xu, J., & Metcalfe, J. (2016). Studying in the region of proximal learning reduces mind wandering. *Memory & Cognition*, *44*(5), 681-695.

Yan, V. X., Bjork, E. L., & Bjork, R. A. (2016). On the difficulty of mending metacognitive illusions: A priori theories, fluency effects, and misattributions of the interleaving benefit. *Journal of Experimental Psychology: General*, *145*(7), 918-933.

Yan, V. X., Soderstrom, N. C., Seneviratna, G. S., Bjork, E. L., & Bjork, R. A. (2017). How should exemplars be sequenced in inductive learning? Empirical evidence versus learners' opinions. *Journal of Experimental Psychology: Applied*, *23*(4), 403-416.

Zulkiply, N., & Burt, J. S. (2013). The exemplar interleaving effect in inductive learning: Moderation by the difficulty of category discriminations. *Memory & Cognition*, *41*(1), 16-27.

Zwaan, R. A., Etz, A., Lucas, R. E., & Donnellan, M. B. (2018). Making replication mainstream. *Behavioral and Brain Sciences, 41*. https://doi.org/10.1017/S0140525X17001972

APPENDIX. EXAMPLE OF BIRD IMAGES USED IN THE CURRENT EXPERIMENT

Note. Top row: Oriole, Warbler, Finch, Swallow. Bottom row: Vireo, Grosbeak, Thrush, Chickadee