FLUENCY AS A CUE FOR ENCODING AND RETRIEVAL

THE IMPACT OF REPETITION ON MEMORY:

FLUENCY AS A CUE FOR ENCODING AND RETRIEVAL

By TAMARA MICHELLE ROSNER, B.Sc. (hons.), M.Sc., M.A.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the
Requirements for the Degree Doctor of Philosophy

McMaster University

DOCTOR OF PHILOSOPHY (2018) McMaster University (Psychology)

TITLE: The impact of repetition on memory: Fluency as a cue for encoding and retrieval

Representations

AUTHOR: Tamara Michelle Rosner, B.Sc. (hons.) (McMaster University), M.Sc.
(McMaster University), M.A. (University of Waterloo)

SUPERVISOR: Professor Bruce Milliken

NUMBER OF PAGES: xx; 183

Lay Abstract

There is an intuitive notion that repeating information increases the likelihood that it will be remembered. Contrary to this intuition, this thesis documents the *repetition decrement effect*: better memory for words seen once than words seen twice in succession. It is proposed that this effect occurs due to signals we automatically use in everyday life to decide if information is worth remembering for later. In general, information we already know is processed more easily than information we do not know. Additionally, repetition makes information easier to process. Therefore, immediately repeating information may make it seem as though you already know it, which may stop you from learning it at all. The repetition decrement effect is then connected to effects of false memory, in which repetition leads to information being incorrectly classified as “known”. These findings speak to the cues we use to navigate the world around us every day.

Abstract

There has long been interest in the effect of repetition on memory amongst cognitive psychologists. A major area of research has examined how repetition at study improves encoding and subsequent memory performance. Another focus in the literature has been on manipulating fluency at retrieval to influence feelings of familiarity, with item repetition at test inducing a classic false recognition effect. Examination of these disparate areas of research hints that similar mechanisms may be operational in producing effects of repetition at study and repetition at test. Work from the false recognition literature suggests that items are more likely to be classified as "old" if they are made to be fluent at test. In other words, fluency may be used as a cue to indicate that information is already known. This fluency attribution process may also influence encoding: if increased fluency signals that information is known, then there may be no need to encode that information. The empirical goal of this thesis was first to better understand the impact of repetition on encoding, and then to better understand the role of fluency when both learning and retrieving information. This thesis documents some of the first examples of a counter-intuitive *repetition decrement effect*, in which items seen a single time are better remembered than items seen twice in succession. Evidence connecting this repetition decrement effect to effects of false recognition is presented, with the suggestion of a common process leading to these two memory effects. More important, this thesis demonstrates the impact of fluency at both encoding and retrieval, and can allow for better understanding of how human cognition operates on a daily basis.

Acknowledgments

Science is an endeavour that cannot be done in isolation. The fact that this thesis exists is due to so, so many people. Whether you helped directly with the research and writing or you provided moral support, thank you, for without you I would be unable to write these words.

The biggest thank you goes to my supervisor, Bruce Milliken. Bruce, I cannot understate how grateful I am to have had the opportunity to work with you. Again. Thank you for your understanding, support, and for keeping me honest when I focused too much on the little details. Thank you for providing guidance on my research, even when you were exhausted after a day of deaning or chairing. Those actions did not go unnoticed. And most important, thank you for giving me the time I needed to figure out that I wanted to come back to the lab and to cognition. Without that generosity, I literally would not be where I am today.

A similar thank you goes to my committee, Scott Watter and Judy Shedden. I truly appreciate you both being so willing to pick up where we left off without any hesitation. Your feedback and support has been invaluable, and I am so glad I was able to finish this journey with the two of you on my team.

To David Shore, my unofficial committee member, thank you for the time and energy you gave to my research. Your input not only made me a better presenter, but a better scientist, and I cannot thank you enough for your guidance.

Thank you to my lab mates for your input on my thesis and for making the office a pleasant place in which to work. To Ellen MacLellan, my travel partner, thank you for your warm and welcoming presence in the lab. You opened your heart and your garage door to me, for which I am forever grateful. To Mitchell LaPointe, thank you for being a patient and calm presence in the office, for being a constant source of support, and for helping me focus on the bigger picture when things felt overwhelming. To Hanae Davis, my no-longer-sprout, your generosity and thoughtfulness are unmatched by anyone I know. Thank you for inspiring me to be a better researcher, scientist, and friend. To Lisa Lorentz, thank you for your willingness to drop everything and listen when I needed to talk. Chatting with you was always a welcome break in the day, whether it was about errors in Python code, silly videos, or life advice. Robert Collins, thank you for your insights into our shared effect, for without them, I would know much less about the finding on which this dissertation is based. To Brett Cochrane, the lab cynic, thank you for your skepticism, for it made my research better.

To past lab members Maria D’Angelo and David Thomson, thank you for your input in my work. Maria, thank you for continuing to be a guiding presence and for your ongoing willingness to teach and advise. Dave, thank you for your brilliant insights into my project and for your Python expertise during your time as a postdoc in the lab.

Of course, no time in Bruce’s lab would be complete without Spanish colleagues. To Juan Lupiáñez, although we’ve only spent a small amount of time together when we run into each other at conferences, talking research with you is always a pleasure. To Raúl López-Benítez, thank you for your collaboration on the first empirical chapter of this thesis. I’m looking forward to grabbing some Tim Hortons together the next time you’re in Canada. To Javier Ortiz-Tudela, thank you for your kindness and generosity, and for inspiring the cookie experiment. We'll do a replication and extension when you visit next.

I also want thank my extended lab family for providing guidance on my research and making the office an enjoyable place to work. To Kaian Unwalla, Connie Imbault, Irina Ghilic, Amy Pachai, Andrew LoGuidice, Brendan Stanley, and Hannah Kearney, thank you for dinners, trivia nights, infamous podcast recommendations, and for simply being there when I needed a chat.

Thank you, too, to the friends I have made in the department. Without you, my graduate school experience would not nearly have been as enjoyable. Thank you for movie nights, BBQs, camping trips, board games, and for making me feel like I never left. You all hold a special place in my life, and even if you think my research is just the Stroop task, I still consider your influence a large driving force in creating this thesis.

To Team Clinical Judgment, thank you for continuing to be there for me, even after I left. My biggest fear was no longer feeling like a part of the team, but that fear could not have been more unfounded. Thank you for the continued nights of Harry Potter, games, Candlenights, and spontaneous get-togethers. And thank you for your support. It truly means the world to me.

To my pre-grad school friends, your continued presence was a welcome constant during an uncertain time in my life. Thank you for always sticking by my side, and for your patience during the last few months when all I had to talk about was writing.

To my family, you are the reason I was able to begin this journey in the first place. Dad, thank you for encouraging my love of science from the beginning. Mom, thank you for getting excited about my work and for being a constant source of support. Adam, thank you for the spontaneous phone chats. They made my day. Savta, thank you for your curiosity about my science, for reminding me that I can continue to learn for a long time to come, and for your patience when you try to talk astrophysics with me and realize that I know less than you. Daphne, thank you for making me who I am today.

Finally, to Evan Borman, thank you for your endless support. Even after all this time, your unflinching belief in my abilities baffles me. I could not have completed this thesis without you, and now that it's finished, I promise I'll have time to play three-hour board games again.

Table of Contents

Preface

This is a “sandwich thesis”, meaning that the empirical chapters are all stand-alone publications that are either published or submitted for publication. Chapter 2 is published in a peer-reviewed journal, and Chapters 3 and 4 are both submitted for publication in peer-reviewed journals. For each of these empirical chapters, I am the first author. For Chapter 2, my collaborators Raúl López-Benítez, Maria D’Angelo, and David Thomson are second, third, and fourth authors, respectively, and my supervisor, Dr. Bruce Milliken, is the final author. Dr. Milliken is the second author of Chapters 3 and 4. My contributions to each of these manuscripts are outlined below.

The first empirical chapter (Chapter 2) is a reprint of Rosner, T. M., López-Benítez, R., D’Angelo, M. C., Thomson, D., & Milliken, B. (2018). Remembering “primed” words: A counter-intuitive effect of repetition on recognition memory. *Canadian Journal of Experimental Psychology, 72*(1), 24-37. doi: 10.1037/CEP0000139. My role in the manuscript included experimental design and programming, data collection from human participants, and data analysis. I was also the primary writer.

The second empirical chapter (Chapter 3) is the following manuscript: Rosner, T. M. & Milliken B. (Submitted). The function of (dis)fluency: A potential cue for encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* Manuscript ID: XLM-2018-0186. My role in the manuscript included experimental design and programming, data collection from human participants, and data analysis. I was also the primary writer.

The third empirical chapter (Chapter 4) is the following manuscript: Rosner, T. M. & Milliken B. (Submitted). The role of expected fluency: A comparison of an old and a new false recognition effect. *Psychonomic Bulletin & Review.* Manuscript ID: PBR-BR-18-124. My role in the manuscript included experimental design and programming, data collection from human participants, and data analysis. I was also the primary writer.

Note that, because these manuscripts are intended to be standalone publications, that there will be some redundancy within the introductions and discussions of these chapters. There is also some redundancy in the methodology across the three chapters, as similar methods were used for the experiments in each chapter. Despite this overlap, each chapter contains unique experiments intended to answer different theoretical questions, all of which are related to the common issues presented in this thesis.

List of Figures

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

List of Tables

CHAPTER 2

CHAPTER 3

CHAPTER 1: Introduction

Human memory is a topic of broad and enduring scientific interest (Ebbinghaus, 1885/1913), but is also a topic that interests the lay public. In general, people are very much interested in how to improve one's memory for information, whether it is for course content in order to do well on a test, the names of people one meets, or just the general want to better remember in everyday life. Moreover, there seems to be an intuition regarding how memory works—both among researchers and the general public—with the sense that repetition must improve memory. This influence of repetition can be observed through typical study strategies (e.g., rote repetition of information), but also through the types of questions researchers ask (e.g., "What *type* of repetition is best for memory?") Contrary to that intuition, this dissertation reports a novel finding in which repetition is *detrimental* to remembering. The goal of this thesis is to better understand the processing that leads to this effect of better memory for non-repeated than repeated items, and connects this effect of encoding to classic false recognition effects in the literature.

To set the context for these goals, I will first provide an overview of recognition memory tasks and the processes thought to drive recognition memory, including the distinction between recollection and familiarity. Then, I will review the topic of familiarity in more detail, describing one possible account of what the subjective feeling of familiarity may reflect. Specifically, I aim to discuss the notion that familiarity arises from an attribution regarding processing fluency, and that processing fluency may indeed provide a valid signal of prior experience. I will then turn to a discussion of false recognition effects and how the fluency attribution process can be used to create false feelings of familiarity, amongst other fluency-based illusions. Finally, I will discuss how repetition has been thought to impact learning and encoding of information, including an overview of the memory strength effect and the spacing effect.

Recognition Memory: Familiarity versus Recollection

Although the focus of the present thesis is not to evaluate the mechanisms that drive recognition memory, it does use recognition memory as tool to study the impact of fluency on encoding and retrieval. As such, the following section will outline how recognition memory is measured and touch on two prominent theories of recognition memory.

In a standard recognition memory task, participants are given items to study (either incidentally or intentionally). They are then are given a memory task in which they are given old and new items and must classify them as “old” or “new”. Performance is typically evaluated by examining “old” responses to old items (hits) and new items (false alarms). The hit minus false alarm difference score is a typical measure of recognition memory (*d’* may also be used, which reflects the difference score after normalizing the hit and false alarm rates; Banks, 1970). Note that “new” responses are typically not examined, as this would provide redundant information. For example, if someone classifies seven out of a possible ten items on a recognition test as “old”, then by default, the final three items must be classified as “new”. Since “new” responses do not provide information that is independent of “old” responses, they are typically not discussed within the recognition memory literature. Although there are variations on recognition memory paradigms (e.g., the use of confidence ratings and remember/know judgments) the basic premise of all methods is the same: asking participants whether or not they recognize an item from earlier in the experiment.

Although the method for collecting recognition memory data is quite simple, understanding what mental processes allow for recognition is decidedly complex, and there is a large debate in the literature as to how recognition memory itself operates. Two broad theories of recognition memory will be touched on here. One of these theories argues that recognition is due to the single process of familiarity (e.g., Banks, 1970; Mickes, Wixted, & Wais, 2007). According to these single-process theories, items in a recognition memory test each have a certain level of familiarity along a continuous scale, as represented in Figure 1. On this scale are two distributions: one that represents the familiarity values of old items and one that represents the familiarity values of new items[1]. On average, old items are higher on this scale (by virtue of having been recently encountered) than new items, though some overlap of these distributions does occur. The familiarity of items on a recognition test is compared to a criterion which is placed along this familiarity scale; items with a familiarity value higher than criterion are classified as "old" and items that have a familiarity value below criterion are classified as "new". According to this view of recognition, sensitivity (the ability to tell old from new items) is based on the separation between the old and new distributions. The more the two distributions overlap, the lower the sensitivity. Moreover, the criterion placement is a

---

[1] For simplicity, Figure 1 represents an equal variance signal detection model. However, it should be noted that proponents of a single process theory typically represent familiarity distributions with unequal variance signal detection models (Mickes et al. 2007).

reflection of response strategy. More liberal strategies will place the criterion lower, meaning that item familiarity does not need to be high for the item to be called "old"; this strategy will ensure that hits are high, but will also drive up false alarms. Alternatively, a more conservative strategy (in which the criterion is high) will ensure low false alarm rates, but also very few hits. Proponents of this theory of recognition memory suggest that all recognition memory performance can be explained by the single process described above, with subjective differences in recognition (e.g., feeling very sure something is old versus being more unsure) reflecting different levels of familiarity strength. In other words, the difference in how well various items are recognized is one of quantity, not quality (see Rotello, Macmillan, Reeder, & Wong, 2005; Wixted & Mickes, 2010 for more complex single-process theories of recognition).

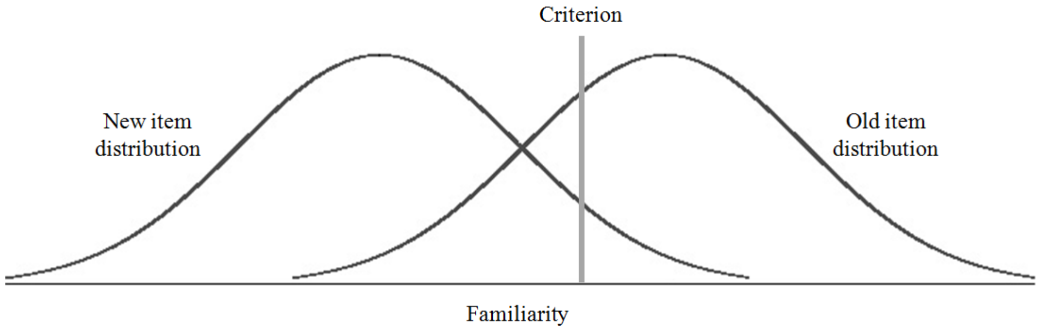

Figure 1: Representation of familiarity along a continuum. The old item distribution is shifted to the right of the new item distribution as old items have greater familiarity values (due to being recently encountered). The criterion is set at a certain familiarity value; items that are higher in familiarity than the criterion will be judged as "old".

Alternatively, dual-process models assume two separate processes are involved in recognition. The model focused on here will be the dual-process signal-detection model (e.g., Yonelinas, 1994), though there are many other dual process models that will not be discussed (e.g., R. Kelley & Wixted, 2001; Whittlesea, 1997). Broadly speaking, dual-process theories also suggest that recognition can occur with a continuous familiarity process, such as the one described above, but they also posit recognition occurs due to a second, all-or-none process of recollection. In other words, recollection either occurs or it does not, rather than recollection strength being considered along a continuum in the same manner as familiarity. The dual-process model discussed here does not assume that recognition occurs via *only* recollection or familiarity, but rather that both can contribute to recognition with varying degrees, and that the level of familiarity for a given item is independent of recollection of that item and vice versa (Parks & Yonelinas, 2007). This view suggests that differences in recognition may be qualitative. For example, just as in single process models, some items may evoke a stronger sense of familiarity than other items, which would be a difference in quantity. However, some items may actually be recalled, with specific details of experiencing that item at study supporting recognition (e.g., "That word reminded me of something my friend said the other day"); this feeling-state is considered a qualitatively different experience from a strong feeling of familiarity that may support performance on a recognition memory task.

As alluded to above, there is an ongoing debate in the recognition memory literature as to whether recognition is due to a single process (e.g., Mulligan & Hirshman, 1995; Rotello et al., 2005; Slotnick & Dodson, 2005; Wixted & Stretch, 2004) or two processes (e.g., Boldini, Russo, & Avons, 2004; Yonelinas, 1997, 2001; Yonelinas, Aly, Wang, & Koen, 2010; Yonelinas, Dobbins, Syzmanski, Dhaliwal, & King, 1996), with each camp using various methods, such as receiver operating characteristics, process dissociation procedures, remember/know procedures, response time analyses, and amnesia studies to support their respective theory (see Wixted, 2007; Yonelinas & Parks, 2007 for reviews). Although the purpose of this thesis, as mentioned, is not to tease apart these theories, it does assume a dual-process theory of recognition. Specifically, an important assumption (particularly in Chapter 3) is that familiarity and recollection are two separate processes, with the former being a fast-acting process and the latter being slower and more controlled (Boldini et al., 2004). Of course, it is possible to reconcile all of the data in the current body of work within a single-process model with the assumption that slower responding during recognition leads to more accumulation of evidence and thus a more accurate familiarity-based recognition response (Mulligan & Hirshman, 1995). However, as the present thesis does not aim to examine this single versus dual process issue in depth, this debate will not be discussed further. Rather, this section is intended to contextualize some key issues in the literature in which the current investigation may be couched. Specifically, understanding the process of familiarity and how it may arise can help facilitate understanding of the empirical work in this thesis, specifically for Chapters 3 and 4. These issues will be discussed in more detail in the next few sections. Next, I will turn to a discussion of the cognitive processes that give rise to the subjective feeling of familiarity in order to better understand what familiarity can tell us about prior experience.

What *is* Familiarity?

Although there is an ongoing debate between proponents of the two outlined theories of recognition memory, most researchers agree that the term "familiarity" is an apt one to describe a process that contributes to recognition memory. Of course, this term simply describes a process, rather than explaining the mechanisms underlying the process, which leads to the question of what exactly *is* familiarity, and why is it experienced? One account in particular will be focused on here, which suggests that feelings of familiarity are not necessarily a direct product of a prior experience being remembered. Rather, familiarity may result from an attribution of processing fluency in the present to a source in the past (Jacoby, Kelley, & Dywan, 1989).

The importance of fluency to recognition memory is particularly apparent in a study conducted by Jacoby and Dallas (1981). In this study, two possible bases for recognition memory performance were explored: the experience of perceptual fluency (ease of processing) and a more controlled search of memory for particular details of a prior encoded item. More specifically, Jacoby and Dallas suggested that one basis for judging an item's oldness is how easily it is processed, with old words generally being processed more easily than new words. A second basis for judging an item's oldness is by how much detail can be generated about the experience of that item in an earlier study phase, with memory for old words but not new words generally being accompanied by such contextual detail. In other words, they explored the possibility of a dual-process theory of recognition memory, though they did not discuss dual-process notions of recognition in the more formal terms outlined above.

The research strategy employed by Jacoby and Dallas (1981) to study recognition memory was particularly informative: recognition performance was examined alongside performance on a perceptual identification task. In a perceptual identification task, participants are presented items (such as words) for a brief duration followed by a pattern mask, and the task is simply to identify the items. Performance on this task is indexed by accuracy. Jacoby and Dallas noted that there were many manipulations that affected performance on both perceptual identification and recognition memory tasks and some manipulations that affected only recognition memory performance. They suggested that when a manipulation affected performance on both tasks, it did so because the performance effects in the two tasks were related; that is, improved recognition memory occurred as a result of the same mechanisms that improved perceptual identification (Jacoby & Dallas, 1981). Specifically, they hypothesized that manipulations at study that enhance perceptual learning would improve performance on both perceptual identification and recognition memory tasks, and that manipulations that improve item elaboration would improve performance on recognition memory tasks only.

Across multiple experiments it was indeed observed that manipulations impacting perceptual learning affected performance on both perceptual identification and recognition memory tasks (Jacoby & Dallas, 1981). For example, in some experiments, repetition was manipulated at study, with some items shown once and some shown twice throughout the study phase. Repetition was thought to be a manipulation that improves perceptual learning as multiple exposures to an item allows for increased learning of perceptual information. In line with that notion, both recognition memory performance and perceptual identification performance were better for items seen twice than items seen once. Similar results were observed when items were shown twice in succession relative to when repetitions were spaced, with better performance on both tasks for spaced than massed items. The impact of modality match, another manipulation thought to impact performance via perceptual learning, was also shown to have an effect on both measures of performance. Across all of these manipulations, perceptual learning was enhanced for one condition over another. This enhanced perceptual learning allowed for more fluent perceptual identification, as well as better recognition memory for those items. Jacoby and Dallas (1981) argued that these two improvements in performance were not independent; rather, they suggested that fluent processing that allows for quick and accurate perceptual identification could also be used as a cue that an item is known. That is, fluency not only enhances item identification, but can also indicate if that item has been previously encountered, thus circumventing the need to rely on more laborious retrieval processes. It is in this way that manipulations that improve perceptual learning at study may improve recognition of those items: not because they are better encoded per se, but because they are more fluently processed at test.

Of course, the impact of perceptual manipulations on identification and recognition was only one part of the Jacoby and Dallas (1981) investigation. As mentioned above, they also aimed to demonstrate that manipulations thought to improve item elaboration at study would result in large improvements in recognition performance without impacting perceptual identification performance. The thought here was that item elaboration at study might allow for the retrieval of contextual information at test, which could be used to support recognition memory but may not be particularly useful in a perceptual identification task. Again, this hypothesis was indeed supported. For example, Jacoby and Dallas (1981) used a levels of processing manipulation in which participants answered questions regarding the letters within presented words (e.g., "Contains the letter L?"; shallow condition) or answered questions about the meaning of words (e.g., "Is the centre of the nervous system?"; deep condition). This manipulation allowed for more item elaboration in the deep condition than the shallow condition while keeping the amount of perceptual learning relatively equal (as participants encountered identical perceptual information between conditions). In line with these predictions, perceptual identification performance was equivalent for the words in the shallow and deep conditions, but recognition performance was better for words in the deep condition than the shallow condition (Jacoby & Dallas, 1981). Similarly, another experiment was conducted in which participants read words or solved anagrams (participants were shown the correct word in the anagram condition after solving each puzzle). Again, the goal of this manipulation was to keep perceptual information equal across conditions, but increase item elaboration for the anagram condition over the word reading condition. Similar to the levels of processing manipulation, there was a large benefit to recognition memory for words that were anagram puzzles over words that were simply read, and there was no difference in perceptual identification between conditions (Jacoby & Dallas, 1981). Yet another manipulation examined was study time, with the idea being that words shown for two seconds can be elaborated on more than words shown for only one second. Again, recognition memory performance was better for items shown for the longer period of time than the shorter period of time, and there was no impact of presentation length on perceptual identification (Jacoby & Dallas, 1981). In general, manipulations that encourage greater item elaboration at study improved recognition performance without impacting perceptual identification performance. These findings demonstrate a qualitatively different process from perceptual learning that supports recognition memory, with the ability to recall specific details from the time of study leading to better recognition memory performance.

Taken together, these findings suggest that we can recognize something as "old" through two different processes. One process occurs if we have had the chance to elaborate on an item when we originally encountered it, which allows for the retrieval of contextual information at test that supports recognition. From the perspective of the dual-process theory of recognition discussed above, this process may be thought of as "recollection". The other process that supports recognition occurs when an item is fluently processed at test, with ease of processing being indicative of a previous encounter. Moreover, this fluency may be subjectively experienced as familiarity. In other words, feelings of familiarity may not reflect the engagement of a retrieval process that supports recognition; rather, familiarity may be a result of an attribution of fluency, making us feel as though something has been previously encountered (Jacoby et al., 1989). Simply put, we can use fluency as a short-cut to recognize: if something is processed easily, we may attribute that easy processing to "oldness", and thus do not need to spend energy recalling specific details to determine if something has been previously encountered.

Therefore, one possible account of what familiarity reflects is an attribution of fluency: if an item is fluently processed, that fluency may be attributed to the item being known, which results in a feeling of familiarity. Indeed, this account not only explains the results of Jacoby and Dallas (1981), but also makes some intuitive sense. Information that has already been encountered is processed more easily (e.g., Jacoby & Dallas, 1981; Witherspoon & Allan, 1985), which means fluent processing can serve as a valid cue that something is known and thus circumvent more effortful retrieval.

Illusions of Familiarity

Of course, fluency can be evoked by many factors, not just prior experience. Fluency has been shown to vary in response to a number of manipulations, such as object masking (e.g., Nairne, 1988), repetition (e.g., Jacoby & Whitehouse, 1989), predictive sentences (e.g., Whittlesea, 1993), and stimulus clarity (e.g., Whittlesea, Jacoby, & Girard, 1990), among many others (see Alter & Oppenheimer, 2009 for a review). Although fluency may be a powerful cue to indicate whether or not something has been experienced previously, fluency is not unique to previously experienced items. As such, illusions of familiarity can be produced in a variety of ways by increasing item fluency at test independent of prior experience with that item.

One instance of this phenomenon is a false recognition effect sometimes referred to as the Jacoby-Whitehouse effect, an illusion of memory that occurs when item repetition is manipulated at the time of recognition (Jacoby & Whitehouse, 1989). The goal of this study was to examine how repetition at test impacted fluency (and thus, recognition memory judgments), with the hypothesis being that repeated words should be more fluent than non-repeated words, and thus produce an illusion of familiarity. After studying a set of words, participants completed a recognition memory test in which target items were preceded by a word that either matched the target (repeated) or mismatched the target (non-repeated). Additionally, participants either experienced long-duration masked primes (200 ms) or short-duration masked primes (50 ms). Participants in the short-duration prime condition produced higher false alarm rates (“old” responses to new items) for the repeated than the non-repeated words. In contrast, participants in the long-duration prime condition produced the opposite effect: more false alarms for the non-repeated than the repeated words (Jacoby & Whitehouse, 1989). Note that these effects were also observed in the hit rates, though they were smaller and less consistent. To explain their findings, Jacoby and Whitehouse (1989) proposed that in the short-duration prime condition participants were unaware of the prime words, so increased fluency due to repetition was attributed to the items being familiar, which increased “old” responses (and false alarm rates) for repeated items. Alternatively, in the long-duration prime condition, it was suggested that participants’ awareness of the primes meant that fluency was appropriately attributed to repetition; this awareness drove down false alarm rates for repeated items, as the fluency experienced due to repetition was discounted when making recognition judgments[2]. Two key principles should be gleaned from these results. First, manipulating repetition at test influences the familiarity of certain words, presumably by

---

[2] Later studies replicating this illusory recognition effect suggest that it is not awareness of the prime per se that may have led to the original Jacoby-Whitehouse effect, but rather the shorter prime duration (Bernstein & Welch, 1991; Joordens & Merikle, 1992). Moreover, other research has since indicated that the Jacoby-Whitehouse effect can be observed with long-duration primes if the salience of repetitions is reduced (Gellatly, Banton, & Woods, 1995) or if attention is drawn away from the primes (Merikle & Joordens, 1997).

changing the fluency with which those words are processed. Second, fluency itself does not equate to familiarity; rather, how that fluency is attributed is important in determining the subjective experience that will arise when encountering fluency. Taken together, these results demonstrate that fluency can be manipulated independent of old/new status and that this fluency can be falsely attributed to an item being “old” under the right circumstances.

Results similar to the Jacoby-Whitehouse effect are observed in experiments that manipulate the clarity of target words (Whittlesea, 1993; Whittlesea et al., 1990). During the recognition test of these experiments, target words were covered with a dynamic visual mask composed of dots (resembling a television channel with poor reception); some of the words were covered with a mask of 20% density, and some were covered with a mask of 40% density. It was generally observed that false alarm rates were higher for words covered with a 20% mask than a 40% mask. That is, the more fluent items produced a greater illusion of familiarity. These results are similar to the Jacoby-Whitehouse effect, demonstrating another manipulation that can be used to increase fluency, and therefore false recognition, independent of old/new status.

The above two examples in which repetition and clarity impacted false alarm rates are generally thought of as instances in which perceptual fluency leads to an illusion of familiarity. However, increased fluency via enhanced conceptual processing can lead to similar false recognition effects (Whittlesea, 1993). In a test phase, words that are preceded by a predictive sentence (e.g., *The shore rocked the BOAT*) typically produce higher false alarm rates than neutral sentences (e.g., *The woman bought a BOAT)*. In this example, perceptual processing of the target word is identical in both cases, and therefore cannot impact recognition judgments. Instead, the manipulation impacts conceptual fluency, or processing of meaning-based information. Conceptual fluency is enhanced for words preceded by predictive sentences, thus producing an illusion of memory similar to those presented above. Moreover, this effect is also observed when participants are asked to identify if a target word is *related* to a word seen at study (rather than an exact match itself), demonstrating that conceptual fluency can produce illusions of various forms of familiarity (Whittlesea, 1993).

All of the studies above indicate that fluency is related (either validly or invalidly) to feelings of familiarity. Again, the key issue to note is that it is the *attribution* of fluency to a source that matters. Moreover, the source to which fluency will be attributed can be manipulated based on the task at hand. For example, it was mentioned previously that manipulating fluency via clarity can impact judgments of oldness; however, the opposite can be true as well. That is, participants will report that a pattern mask covering old items is less dense than a pattern mask covering new items, even if the noise is the same in both cases (Whittlesea et al., 1990). In other words, an illusion of clarity can be caused by increased fluency for old items. A similar effect has been observed for duration judgments, with old words being judged as shown for longer periods of time than new words, even if both are shown for the same short duration (Witherspoon & Allan, 1985). Illusions of duration can also be produced by varying item clarity, with items shown with a less dense mask judged as being presented for a longer period of time than items with a denser mask (Whittlesea, 1993). A host of other manipulations of fluency can influence behaviour in many other tasks. For example, easily pronounced fake stock names are estimated to be more successful in the market than fake names that are harder to pronounce (Alter & Oppenheimer, 2006); conceptual fluency can make a piece of information seem more true or accurate (C. M. Kelley & Lindsay, 1993); greater stimulus-to-background contrast leads to stimuli being judged as more likeable (Reber, Winkielman, & Schwarz, 1998); previously encountered names are more likely to be judged incorrectly as famous than new names (Jacoby, Woloshyn, & Kelley, 1989); and aphorisms that rhyme are judged as more true than those that do not rhyme (McGlone & Tofighbakhsh, 2000). These examples demonstrate that numerous manipulations of fluency can influence behaviour on a wide variety of tasks (see Alter & Oppenheimer, 2009 for a review). The key thing to note in all of these cases is that fluency is typically attributed to the property that is task-relevant; that is, fluency based on oldness will be attributed to clarity if clarity is being judged, and fluency based on clarity will be attributed to oldness if oldness is being judged. Overall, it is apparent that our ability to discriminate between sources of fluency is poor, and we instead attribute fluency to the most obvious source available to us (Whittlesea, 1993; see also Olds & Westerman, 2012).

Absolute Versus Relative Processing Fluency

The above summary highlights the idea that processing fluency may lead to feelings of familiarity if someone is being asked to judge the old/new status of an item. However, the precise processes by which fluency is attributed to familiarity are themselves an important topic of debate. In particular, whereas it may seem reasonable to conclude that it is high absolute levels of processing fluency that lead to attributions of familiarity, Jacoby and Dallas (1981) noted that it is *relative* fluency of an item, rather than absolute fluency, that is key to fluency-based feelings of familiarity. For example, Westerman (2008) found that fluency relative to global context is important in observing a repetition-based false recognition effect. The experiment conducted was based on the method used by Jacoby and Whitehouse (1989), with a test phase consisting of words that were preceded by quickly shown primes that either matched or mismatched the target. Critically, Westerman manipulated the proportion of repeated and non-repeated words between four different groups of participants, with 10%, 33%, 67%, or 90% of the items being repeated. If absolute fluency (i.e., item fluency independent of context) is responsible for producing a false recognition effect, then repeated items should produce higher false alarm rates than non-repeated items across all four groups, and the size of the effect should not differ. If instead relative fluency is important, then the false recognition effect should be largest when fluency is rare and reduce in size as fluency becomes more common within the global context of the experiment. Sure enough, the classic Jacoby-Whitehouse effect was largest for the group in which only 10% of the items were repeated, and the effect disappeared for the group in which 90% of the items were repeated (Westerman, 2008). Similar results were observed in an experiment in which the proportion of predictive sentences was manipulated, again showing that the false recognition effect was largest in the condition in which predictive sentences were rare (Westerman, 2008). These results demonstrate that relative fluency, not absolute fluency, may lead to feelings of familiarity, with rare fluency leading to larger false recognition effects.

The importance of relative processing fluency (Jacoby & Dallas, 1981) can also be observed in studies that contrast expected fluency with actual fluency (Whittlesea & Williams, 1998, 2000, 2001). Whittlesea and Williams (1998) noted that attributions of familiarity depend on a comparison of actual fluency relative to expected fluency in a given situation. The striking example they cited was the experience of seeing your spouse in the kitchen. Although your spouse's face will be processed with a high level of fluency—after all, they are someone you know well and see often—you typically do not experience a strong sense of familiarity in this situation. However, if you were to encounter your spouse at your place of work (an unexpected context for this person), a strong sense of familiarity is likely to be evoked. Simply put, seeing a "familiar" person only leads to a feeling of familiarity when they are seen in a surprising context (Whittlesea & Williams, 1998). To explain this phenomenon, Whittlesea and Williams (1998) put forward the discrepancy-attribution account for fluency-based illusions (see also Whittlesea & Williams, 2000, 2001). This account suggests that there is on-line development of expectations regarding how a stimulus will be processed, and this expectation is compared to actual processing. If there is a discrepancy between expectation and reality, then that *surprising* level of fluency will then be attributed to the most likely source available (such as familiarity).

To demonstrate this idea, Whittlesea and Williams (1998, 2000, 2001) conducted a series of studies using regular non-words: words that have no meaning, but follow orthographic rules and very much seem like words that could be in the English lexicon (e.g., *HENSION, FRAMBLE*). Across experiments, they demonstrated that these regular non-words are less fluent than actual words (as indexed by pronunciation latency), and yet these regular non-words produced higher false alarm rates than actual words; that is, the *less* fluent item produced a *greater* illusory recognition effect. This result can easily be explained by the discrepancy-attribution account. When non-words are encountered, there is an expectation that they will be disfluent since they are not a common occurrence in everyday life. However, these non-words are designed to be similar to actual words, and due to their orthographic structure, are high in fluency. Therefore, the expectation for these non-words (disfluent) and the actual experience they produce (fluent) do not align, which produces surprise. In a recognition task, this discrepancy between expectation and outcome is attributed to the item having been previously experienced. This attribution is consciously experienced as a feeling of familiarity, and thus leads to increased "old" responses (Whittlesea & Williams, 1998). On the other hand, when actual words are encountered, they are expected to be fluent, and are indeed fluent. Therefore, the expectation for actual words (fluent) and the experienced outcome (fluent) do align, so the fluent processing is not a surprise. As there is no discrepancy between expectation and outcome the attribution of fluency is not necessary and no feeling of familiarity is evoked.

It should be noted that this discrepancy-attribution account can explain other fluency-based illusions and not just false recognition effects. For example, Whittlesea and Williams (1998) used this account to demonstrate an illusion of duration, in which surprising fluency led to the perception of words being shown for longer periods of time, independent of presentation duration. Again, the *attribution* of surprising fluency, and not the surprising fluency itself, determines what type of illusion will be produced; the surprise produced by processing discrepancy is attributed to the most likely source in a given context (e.g., oldness in a recognition test; presentation time in a duration judgement task).

Taken together, these findings demonstrate that fluency can be manipulated in a variety of ways, and how we attribute that fluency can also be manipulated such that illusions of recognition can easily be evoked. Although fluency may generally be a good signal to determine oldness (thus producing feelings of familiarity), this signal is not perfect. As such, false recognition effects can be produced in many ways and in many circumstances simply by manipulating fluency independent of old/new status.

As highlighted above, the impact of repetition on illusory recognition has long been used to demonstrate how fluency can lead to feelings of familiarity. However, repetition has not only been used to study false memory. A large area of interest in the memory and education literature concerns how repetition at study impacts the acquisition of information. Specifically, researchers have investigated the conditions under which repetition produces the greatest benefit to learning. This is the topic I will turn to next, focusing on a discussion of how stimulus repetition at the time of study has thought to benefit encoding. Following this discussion, I will connect the literature on stimulus repetition and memory to the false recognition literature in order to better contextualize the goals of the present thesis.

Stimulus Repetition and Encoding

The influence of stimulus repetition during encoding has been a target of many studies of memory, with a particular interest in how item encoding improves under various repetition conditions. For example, many studies have looked at what may be referred to as the memory strength effect: the finding that items shown multiple times in a study phase produce better performance on a recognition memory task than items shown once (Bruno, Higham, & Perfect, 2009; Jacoby & Dallas, 1981; Stretch & Wixted, 1998; Yonelinas, 1994). This finding is a robust one, and has been used to study the processes involved in recognition memory and inform theories of recognition memory. For example, as mentioned previously, Jacoby and Dallas (1981) demonstrated that items shown twice in a study phase are more easily identified and better recognized than items shown once, suggesting that the extra perceptual learning afforded to repeated items leads to increased perceptual fluency, which is then used as a cue for oldness. Other researchers have used the memory strength effect to examine various versions of the single-process model (Bruno et al., 2009; Stretch & Wixted, 1998). Although there may be some debate over what *causes* the effect, the effect itself is clear and replicable: when repetitions are spaced throughout encoding, items shown multiple times are better remembered than items shown once.

The effects of repetition have also been shown to be most beneficial when repetitions are spaced than when repetitions are massed (e.g., Baddeley & Longman, 1978; Bahrick, 1979; Bahrick & Phelps, 1987; R. A. Bjork & Allen, 1970; Braun & Rubin, 1998; Cuddy & Jacoby, 1982; Hintzman, 1974; Jacoby, 1978; Kahana & Howard, 2005; Underwood, 1969). This ubiquitous finding known as the spacing effect (see E. L. Bjork & Bjork, 2011; R. A. Bjork, 1994; Hintzman, 1974) demonstrates that not all repetitions are equal, and that learning is best served by spacing repetitions. For example, some studies have shown that spacing study sessions over many days leads to better final performance than having an equal amount of study time in a single session (Baddeley & Longman, 1978; Bahrick, 1979; R. A. Bjork, 1994). Within these experiments, participants are given information to learn (such as definitions of Spanish words or a motor skill), and are asked to do all of their learning within a single long session (massed), or many short sessions (spaced). Within the study session(s), participants in the massed condition typically show better acquisition than participants in the spaced condition, with those in the spaced condition needing to re-study information more often when trying to learn than those in the massed condition. However, on a final test, those in the spaced condition perform better than those in the massed condition, demonstrating better long-term retention when tested months or even years later (Baddeley & Longman, 1978; Bahrick, 1979; Bahrick & Phelps, 1987).

The spacing effect has been shown to be operative even within single study sessions (Braun & Rubin, 1998; Cuddy & Jacoby, 1982; Jacoby, 1978; Kahana & Howard, 2005; Underwood, 1969). Typically, participants are given a study phase in which they are presented items twice; some items are repeated immediately (massed) while some items are repeated with a variable number of other items between them (spaced). It has been observed that even this simple manipulation can produce better memory for spaced than massed items for many different memory tasks, such as recognition (Braun & Rubin, 1998), free recall (Braun & Rubin, 1998; Kahana & Howard, 2005; Underwood, 1969), and cued recall (Cuddy & Jacoby, 1982; Jacoby, 1978). It is suggested that this may be due to extra learning that takes place for the second occurrence of an item when the repetitions are spaced, and that this extra learning is absent when the repetitions are massed. More specifically, this version of the spacing effect may occur because in a massed condition, one can simply "recall" the item they had just seen rather than having to work to identify the item a second time. The study conducted by Cuddy and Jacoby (1982; see also Jacoby, 1978) more concretely demonstrates the distinction between "solving a problem versus remembering a solution" (Jacoby, 1978, p. 649). Participants in their experiments completed a study phase in which related words (e.g., TREE: BRANCH) were presented. All pairs were presented twice throughout the study phase and were either massed or spaced. Critically, for half of the pairs the second word was presented with missing letters for its second presentation (e.g., TREE: BR--CH; read-construct pairs), whereas the other half of pairs had both words presented intact (read-read pairs). The participants' task during the study phase was to read both words for the read-read pairs and to read the first word and then complete the second word for the read-construct pairs. Therefore, the correct response for read-read pairs was always readily available to participants regardless of spacing, as they simply had to read the words presented. For the massed read-construct pairs the solution was also readily available (as the solution had just been read on the prior trial), whereas for the spaced condition participants had to work a bit harder to determine the solution. As such, the most learning could be said to occur for the spaced read-construct pairs.

Following the study phase, participants were given a surprise cued-recall test in which the first word in a pair was presented and they had to provide the second word. Performance was better for the read-construct pairs than the read-read pairs, but only when presentations were spaced (Cuddy & Jacoby, 1982). When extra effort was required in the spaced condition to complete the read-construct pairs, the second word was better encoded than for read-read pairs and for massed read-construct pairs. These results demonstrate a key finding from the spacing effect literature: spacing study opportunities may make learning more difficult at one point in time but can greatly benefit performance later on.

Taken together, it seems that repetition is generally beneficial to recognition (as evidenced by the memory strength effect), but that immediate repetition reduces that benefit (as evidenced by the spacing effect). Therefore, there are processes inherent to encoding massed repetitions that impede effective learning. Although it has been suggested that the spacing effect is due to better learning opportunities for spaced than massed items, part of this effect may also be related to more fluent processing of massed than spaced items. It has been well-documented that repetition of an item can lead to increased processing fluency, particularly for immediate repetitions (e.g., Jacoby & Whitehouse, 1989). Therefore, it seems plausible that massed repetitions in the spacing effect literature are generally more fluently processed than spaced repetitions (also as seen in Cuddy & Jacoby, 1982). Following this logic, it may be the case that the spacing effect occurs due to increased fluency for massed repetitions being misattributed to those items already being "known" (similar to the attribution process in the false recognition literature), leading to poor learning of that item—after all, why remember something you already know? Overall, the spacing effect may point to fluency as a signal for whether or not resources should be dedicated to encoding an item.

A key point that should be emphasized is that, throughout this dissertation, encoding is assumed to occur regardless of intent to remember. That is, even in experiments in which participants are unaware of an upcoming memory test, encoding processes are assumed to be operative. In other words, encoding is viewed as a process that allows us to learn information about the world around us, to more efficiently respond to and interact with the world, rather than exclusively to remember the details of our prior experiences. This view is supported by object file (Kahneman, Treisman, & Gibbs, 1992) and event file (Hommel, 1998, 2004) theories of perception, which state that responses to stimuli/situations/events can be facilitated by recalling information about previously encountered and similar stimuli/situations/events. In other words, encoding information at one point in time can help us with responding to similar information later. Here and throughout this dissertation, I assume that encoding processes are always engaged regardless of conscious intent, and more important, that our cognitive systems select information that may be most useful to encode.

With this assumption stated, it may be the case that fluency is used as a cue to discriminate information that should be encoded from information that should not be encoded. Put simply, if a particular event is fluently processed, then this fluency may signal that the event is well-represented in memory, and learning and encoding additional information will be of little use. In contrast, situations that are less familiar and more complex may be important to learn, so that the next time a similar situation is encountered it can be dealt with more efficiently.

Furthermore, this principle of difficulty leading to better learning (sometimes referred to as desirable difficulty; E. L. Bjork & Bjork, 2011; R. A. Bjork, 1994) is not just observed in the spacing effect literature. Across many experiments, better memory has been observed for stimuli and information that is more difficult to process, with better recognition performance for masked than intact words (Hirshman & Mulligan, 1991; Hirshman, Trembath, & Mulligan, 1994; Mulligan, 1996; Nairne, 1988); for more disfluent than fluent fonts (Diemand-Yauman, Oppenheimer, & Vaughan, 2011; Oppenheimer & Frank, 2008); for items that are generated than items that are read (Slamecka & Graf, 1978); for low frequency than high frequency words (Glanzer & Adams, 1985; Gregg, 1976); for incongruent than congruent stimuli (Krebs, Boehler, De Belder, & Egner, 2015; Rosner, D'Angelo, MacLellan, & Milliken, 2015); and for blurry than clear stimuli (Rosner, Davis, & Milliken, 2015; but see Hirshman et al., 1994; Yue, Castel, & Bjork, 2013). Moreover, one theory suggests that these effects are due to disfluent information encouraging item-specific processing and thus leading to superior recognition performance, supporting the notion that difficult-to-process information may initiate encoding mechanisms that aim to better understand that information (whereas fluent items may encourage more broad relational encoding that is less focused on the item itself; McDaniel & Bugg, 2008). Relating these theories back to the spacing effect, we can assume the second instances of massed items are more fluent than the second instances of spaced items. Therefore, the second occurrences of spaced items may be better encoded than the second occurrences of massed items since the difference in fluency may be a signal that spaced items need to be better learned. This additional learning in turn would lead to the finding of better memory for spaced than massed items.

To summarize, the benefits of repetition at encoding appear to be greater when repetitions are spaced than when they are massed. Moreover, both theory (e.g., R. A. Bjork, 1994; McDaniel & Bugg, 2008) and data (e.g., Cuddy & Jacoby, 1982; Jacoby, 1978) suggest that this finding may be due to more difficult or disfluent processing of the second instance of spaced pairs than massed pairs leading to better learning. This idea points to a possible connection between the spacing effect (and effects of repetition at encoding more generally) and repetition-driven false recognition effects in which immediate repetition and increased fluency impacts both encoding processes and feelings of familiarity.

Overview of Empirical Chapters

Although there has been much interest in the literature regarding how massed versus spaced repetitions impact memory and learning, researchers have not yet examined how immediate repetitions compare to immediate alternations. That is, it remains an open question as to whether or not an immediate repetition results in better memory for that information than only encountering the information a single time. Previously, it was suggested that fluency may be used as a signal to determine if something is already known. Moreover, if an item is determined to be known, then that item may not need to be encoded. Therefore, immediate item repetition that increases item fluency may result in poor encoding, and comparing immediate repetition to a single instance may produce a counter-intuitive result: better memory for items seen once than items seen twice in succession. Of course, the opposite result may be found in which repetition leads to better memory than a single exposure, simply because repeating an item allows for an extra opportunity to learn that item.

To that end, the goal of this thesis is to more carefully examine the impact of repetition on memory performance (both at the time of encoding and retrieval). Specifically, in Chapter 2 of this thesis, I examine the effect of immediate repetition on encoding and use items presented once (non-repeated) as a comparison point. In Chapter 3, I connect the finding of better memory for non-repeated than repeated items to the classic false recognition effect described by Jacoby and Whitehouse (1989) to better understand the processes resulting in this curious memory effect. Finally, in Chapter 4, I aim to more carefully examine the effect of repetition on false recognition by comparing the impact of repetition during retrieval to another manipulation that results in a similar illusion of familiarity. All of this work points to how fluency attributions can impact memory at various stages of remembering, and I argue that similar processes may influence both information encoding and retrieval.

CHAPTER 2: Remembering "primed" words: A counter-intuitive effect of repetition on recognition memory

Rosner, T. M., López-Benítez, R., D'Angelo, M. C., Thomson, D., & Milliken, B. (2018).
Canadian Journal of Experimental Psychology, 72(1), 24-37.
doi: 10.1037/CEP0000139

Preface

Chapter 2 presents the results of four experiments in which the effect of immediate repetition on memory was examined. The general method for all of the experiments involved presenting green prime words followed by red target words to participants during an incidental study phase. Participants were told to read the red target word aloud at study. At test, they were to classify red words as old or new. In Experiment 1, it was observed that repeated words were remembered less well than non-repeated words on a following recognition memory test. Experiment 2 demonstrated that this effect of repetition on memory could not be attributed to source memory for word colour and Experiment 3 showed that this effect could not be disrupted with short-lag spacing. Finally, Experiment 4 managed to reverse this effect of repetition on memory by having participants read aloud both the prime and target words, producing better memory for repeated than non-repeated words. The results of these experiments show that repetition may impair encoding, with the suggestion that this effect may be due to reduced attention toward repeated words and/or increased attention toward novel (non-repeated) words. These results constitute a first demonstration of the repetition decrement effect (though the effect had not been named as such within this publication).

Abstract

The present study examines the effect of immediate repetition on recognition memory. In a series of 4 experiments, the study phase task was to name aloud a word that was immediately preceded by either the same word (repeated trials) or a different word (not-repeated trials). Across experiments, performance in the study phase demonstrated the anticipated benefit in naming times for repeated trials. More important, performance in the test phase revealed greater sensitivity for not-repeated than repeated trials. This effect was observed even when repetitions at study were separated by an unrelated word (Experiment 3), and was eliminated only when participants named both words in succession at study (Experiment 4). These findings fit nicely with the desirable difficulty principle (R. A. Bjork, 1994), as they demonstrate that items more easily processed at study (i.e., repeated items) are not as well-encoded as items that are more difficult to process at study (i.e., not-repeated items). Furthermore, the current study points to the possibility that attentional orienting in response to processing difficulty may constitute a broadly important cognitive control adaptation that impacts memory encoding.

Introduction

Experimental psychologists have used stimulus repetition in diverse ways to understand basic principles of perception, learning, and memory. The present study focuses on a tension between two opposing influences of stimulus repetition. On the one hand, it has long been known that stimulus repetition facilitates perceptual identification (e.g., Jacoby & Dallas, 1981; Scarborough, Cortese, & Scarborough, 1977), and that multiple opportunities to encode the same stimulus lead to improved retention (e.g., R. A. Bjork & Allen, 1970; Cuddy & Jacoby, 1982). On the other hand, stimulus repetition impedes the orienting of attention; attention tends to be captured by new rather than old perceptual objects (Yantis & Jonides, 1984), and it shifts referentially to locations and objects that have not recently been attended (Posner & Cohen, 1984).

The tension between these two well-documented influences of stimulus repetition rests in that they appear not to fit with a third well-documented principle: attention is fundamentally important to memory encoding (e.g., Craik, Govoni, Naveh-Benjamin, & Anderson, 1996). If stimulus repetition impedes attention, and if attention is fundamental to memory encoding, then it seems that stimulus repetition must hurt memory encoding. Yet this prediction runs counter to intuition—there is a general sense that remembering typically benefits from multiple opportunities to engage in memory encoding. The present study examines the relation between immediate stimulus repetition (i.e., repetition over a very short temporal interval) and memory performance. To our knowledge, no study to date has demonstrated that immediate stimulus repetition can hurt memory performance. In the present study, we describe several experiments in which this result does indeed occur—superior memory performance for a word presented once is better than for a word presented twice in rapid succession.

This result was discovered as part of a research program that examined the link between cognitive control and recognition memory. Our first study in this line of research tested the influence of selective attention at the time of encoding on recognition memory (Rosner, D'Angelo, MacLellan, & Milliken, 2015), whereas a subsequent line of research examined the influence of perceptual degradation on recognition memory (Rosner, Davis, & Milliken, 2015). In these studies, we found that items that were difficult to process at study (i.e., incongruent rather than congruent selective attention items, or blurry rather than clear items) were recognised more accurately at test. These results led us to examine whether a related result might occur with a manipulation of stimulus repetition at the time of encoding, and we report here that indeed it did. Although not originally motivated by this idea, we came to appreciate that the results of these studies are a good fit for the desirable difficulty principle introduced long ago by R. A. Bjork (1994).

The Desirable Difficulty Principle

Desirable difficulties are defined as conditions that increase encoding difficulty, but in doing so improve learning (E. L. Bjork & Bjork, 2011; R. A. Bjork, 1994). For example, completing practice tests is typically more difficult than repeated studying of to-be-learned information, but leads to better performance on a final test (Roediger & Karpicke, 2006). Varying the conditions associated with encoding, and intermixing rather than blocking encoding conditions, also increase encoding difficulty but lead to better retention. Perhaps the best example of desirable difficulty is the spacing effect. In spacing effect studies, information is studied (or learned) in either a massed condition (in which study sessions occur in close succession) or a spaced condition (in which study sessions are separated by longer amounts of time). Learning typically occurs more slowly in the spaced than massed condition, but performance on a final recall test is superior in the spaced condition. Moreover, the longer the interval between study sessions, the better the final performance (Baddeley & Longman, 1978; R. A. Bjork & Allen, 1970)

A nice example of the spacing effect was reported by Bahrick (1979). In a first study session, individuals were presented with English–Spanish word pairs. In subsequent study sessions, participants were presented with English words and asked to provide the Spanish translation. The study sessions continued until all words were correctly translated. Critically, study sessions were separated by an interval of 0, 1, or 30 days. Those in the 0-day interval condition completed all study sessions in one sitting. Performance in the study sessions was best for short intersession intervals; that is, the shorter the interval, the better individuals were at recalling the Spanish translation. However, in a final retention test that was administered months after the study sessions, individuals in the spaced conditions (1 or 30 days) performed better than those in the massed condition. Moreover, recall was better in the 30-day than the 1-day condition. Even more interesting, a follow up study conducted 8 years later revealed that individuals in the spaced conditions performed better on a test of the learned words than those in the massed condition (Bahrick & Phelps, 1987).

The spacing effect can also be observed with short-term spacing. That is, within a single session, words that are shown in massed presentations (i.e., two successive presentations) are not as well remembered in a following memory test as those in spaced presentations (i.e., with other items separating the two presentations). Again, greater spacing between items at the time of encoding tends to produce larger memory benefits for repeated items (Braun & Rubin, 1998; Cuddy & Jacoby, 1982; Hintzman, 1974; Jacoby, 1978; Kahana & Howard, 2005; Underwood, 1969).

Perceptual Desirable Difficulties

Desirable difficulties are also evident in studies that manipulate perceptual characteristics to alter the ease of identification. For example, words presented in unfamiliar fonts are identified more slowly at encoding but remembered better than words presented in familiar fonts in both laboratory and classroom settings (Diemand-Yauman, Oppenheimer, & Vaughan, 2011). The perceptual interference effect may also constitute an example of a perceptually based desirable difficulty (Hirshman & Mulligan, 1991; Hirshman, Trembath, & Mulligan, 1994; Mulligan, 1999; Nairne, 1988). In studies of this effect, words are presented briefly at study either with or without a following pattern mask. Of course, masked words are more difficult to identify than unmasked words at study but are often better remembered during subsequent memory tests.

As noted above, the present study followed from two earlier studies in our lab, both of which produced effects that might be described as perceptual desirable difficulties. In one study, Rosner, Davis, et al. (2015) conducted a series of experiments in which words were presented in a clear or blurry font at the time of study. The study phase task was simply to name the words aloud. Naming times were slower for blurry than clear words. However, the blurry words were better recognized than clear words at the time of test (see also Rosner, Davis, et al., 2015, and Yue, Catsel, & Bjork, 2013, for limiting conditions of this effect). In another study, item congruency was manipulated using a selective attention method at the time of encoding (Rosner, D'Angelo, et al., 2015). In an incidental study phase, a red and a green word were interleaved and presented to participants, with the task being to read the red word aloud. The two words were either the same (congruent) or different (incongruent). Naming times were about 100 ms slower for incongruent than congruent items, but incongruent items were recognized more accurately than congruent items on a surprise recognition memory test (see Krebs, Boehler, De Belder, & Egner, 2015, for a similar result in a gender identification task).

The Present Study

The goal of the present study was to examine whether perceptual desirable difficulty effects might be produced by manipulating immediate stimulus repetition at the time of study. The method used by Rosner, D'Angelo, et al. (2015) to study congruency effects on recognition required only a slight adaptation to study this issue. In that study, a red word and a green word were presented simultaneously, and participants were required to name just the red word. In the present study, we modified this procedure simply by offsetting the red word and green word in time; that is, a green word prime was followed shortly after by a red word target, with the task again being to name the red word target. Based on the results observed by Rosner, D'Angelo, et al. (2015), we predicted better memory performance for not-repeated items than for repeated items.

The prediction that stimulus repetition might hurt memory performance is somewhat counterintuitive; one might expect repeated words to be better remembered than not-repeated words simply by virtue of being seen more often. On the other hand, the idea that stimulus repetition might hurt memory performance is in line with the perceptual desirable difficulties observed in our prior studies, as well as the observation that the orienting of attention is impeded when stimuli are repeated (Klein, 2000; Posner & Cohen, 1984; Spadaro, He, & Milliken, 2012; Yantis & Jonides, 1984). Moreover, immediate stimulus repetition can be thought of as an extreme instance of massed study. As seen in the spacing effect literature, massed presentation of items impedes learning, and it seems conceivable that the mechanism that underlies poor memory for massed presentation might also contribute to relatively poor memory for immediate repetitions at study.

Experiment 1

The purpose of Experiment 1 was to examine the effect of stimulus repetition on later recognition using an adaptation of the procedure used by Rosner, D'Angelo, et al. (2015). Rather than presenting red and green interleaved words simultaneously, a single green word prime was followed by a single red word target in the study phase. The study phase task was simply to read the red word aloud. At the time of test, the target words from the study phase were intermixed with new words and participants were required to make old/new recognition judgments. Naming times in the study phase were expected to be faster for repeated than not-repeated items. The key issue here is whether recognition would be better for the not-repeated than repeated items as predicted above.

An additional purpose of Experiment 1 was to replicate the aforementioned effect of congruency on recognition memory. As such, Experiment 1 included a second group that completed a direct replication of the selective attention method of Rosner, D'Angelo, et al. (2015). Therefore, participants were assigned to one of two groups: the congruency group (the replication of Rosner, D'Angelo, et al., 2015) or the repetition group. The results from the congruency group replicated the pattern observed in our prior study; recognition was better for incongruent than for congruent items. However, as the results from this group were not directly relevant to the present study of repetition effects on recognition, they are not presented in detail here. The interested reader will find details of the method and results from the congruency group in the supplementary materials section included with this manuscript.

Method

Participants. Twenty-four participants (18 female, mean age = 18, *SD* = 0.99) were recruited from the McMaster University student pool in exchange for course credit. All participants spoke English fluently and had normal or corrected-to-normal vision. Written consent was obtained from all participants prior to conducting this and all subsequent experiments.

Apparatus. The experimental procedure was run using PsychoPy (Peirce, 2007, 2009) on a Dell computer. Stimuli were presented on a 24 in. BENQ LED monitor.

Stimuli. All stimuli were red and green words presented on a black background and were presented in Lucida sans console font at the center of the screen. The words were five-letter, high frequency nouns (Thorndike & Lorge, 1944) and were presented with a space between each letter. Each word subtended 5.95 degrees of visual angle horizontally, and 0.75 degrees vertically. The exact positioning was based on the interleaved stimuli used for the congruency group (see supplementary materials for details), so that the words were slightly offset from one another both vertically and horizontally. This offset was subtle and unlikely apparent to participants, as the red and green words were never on screen at the same time. Examples of the stimuli are presented in Figure 1.

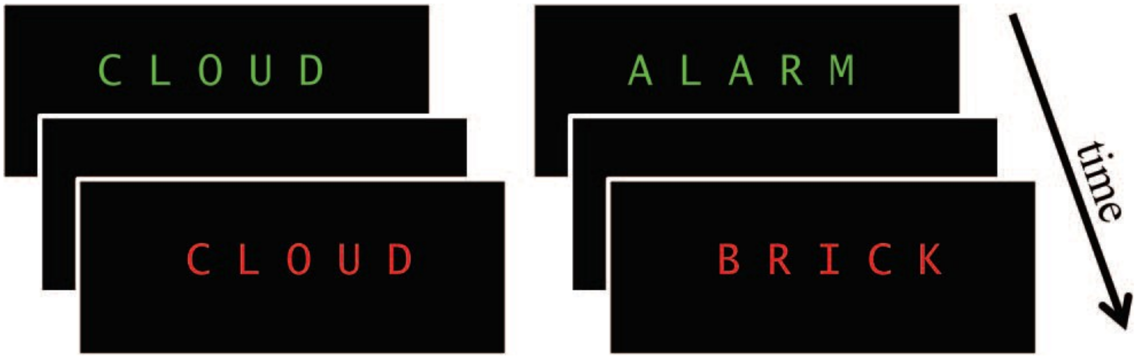

Figure 1. Examples of study phase items for Experiment 1. The left side depicts an example of a repeated item; the right side depicts an example of a not-repeated item. See the online article for the color version of this figure.

Procedure. The experiment took place in a well-lit room with participants tested individually. Participants sat 50 cm in front of the monitor, with a standing microphone placed in front of them. This microphone recorded participants' response times for each trial during the study phase from the time of trial onset to the time of voice onset. In the study phase, participants were told that they would see a green word followed by a red word on each trial, and they were to say the red word aloud as quickly and accurately as possible. They were not informed of a later recognition memory test. Each trial began with the presentation of a white fixation cross in the center of the screen for 2,000 ms, followed by a green prime word for 500 ms, a 250 ms blank interstimulus interval (ISI), and a red target word for 1,000 ms. For half of the trials, the green prime and red target had the same identity (repeated trials). For the other half of the trials, the green prime and red target had different identities (not-repeated trials). Following each naming response, the experimenter coded participants' responses as correct, incorrect, or a spoil by pressing "1", "2", or "3" on the keyboard, respectively. Responses were classified as incorrect if participants said the wrong word, said the wrong word and then corrected their response, or failed to provide a response. Responses were classified as a spoil if participants made a spurious response (e.g., stuttering, throat clearing) that was thought to trigger the microphone early.

Following the study phase, participants completed a 10-min distractor math task to serve as a filled retention interval, after which they were given instructions for the surprise recognition memory test. Participants were told that they would see a red word on every trial, and they were to indicate whether or not they remembered reading the word during the previous study phase. If they remembered the word, they were to respond "old" by pressing the A key on the keyboard; otherwise, they were to respond "new" by pressing the L key.

The recognition memory test also included a remember/know distinction for items given an "old" response (Rajaram, 1993). However, the labels "Type A" and "Type B" were used in lieu of "remember" and "know," respectively, as doing so reduces the overall level of false alarms (McCabe & Geraci, 2009). Participants were told that if an item was classified as “old”, they would then be asked to specify if the item provoked a Type A memory (a feeling of remembering) or a Type B memory (a feeling of knowing). After receiving instructions, participants were asked to describe the difference between Type A and Type B memories to the experimenter in their own words, to ensure they understood the distinction.

Every trial for the test phase began with a white central fixation cross for 2,000 ms, followed by a red word and the words “OLD” and “NEW” on the bottom left and right of the screen, respectively, to serve as reminders for which key corresponded to which response; these stimuli stayed on screen until response. When participants made a “new” response, the next trial began immediately after. When participants made an “old” response, the words “OLD” and “NEW” were replaced by “TYPE A” and “TYPE B”, once again to serve as reminders of the remember/know distinction that was to be made. These stimuli stayed on screen until response, at which point the next trial was initiated.

Design. The stimuli were constructed using six lists of 60 words (see Appendix A). Three of the word lists were assigned to the “old” stimulus set, and the other three were assigned to the “new” stimulus set. These roles were counterbalanced across participants. Within the old stimulus set, one list was used to construct 60 repeated items, with each item in the list used as both a prime and a target. The other two lists were used to construct the not-repeated items, one list for primes and the other for targets. The assignment of lists to these three roles for old items was counterbalanced across participants. Within the new stimulus set, the three lists were randomly assigned across participants to serve as repeated items, as not-repeated primes, or as not-repeated targets.

It should be noted that, because only red target words were presented at test, the word list assigned to the “new not-repeated prime” role for each participant was coded as such, but never actually seen during the experiment. The slight horizontal and vertical offset of prime word relative to target word was counterbalanced across items. There were 120 trials in the study phase (60 repeated and 60 not-repeated items) and 240 trials in the test phase (120 old items and 120 new items). All targets from the study phase were presented during the test phase. The order of presentation of items was randomized in both the study and test phases, with repeated and not-repeated items intermixed during the study phase, and old repeated, old not-repeated, and new items intermixed during the test phase.

Results

Correct naming times in the study phase and recognition sensitivity in the test phase were the key dependent variables here and in all subsequent analyses in this article.

Study phase. Correct response times (RTs) were first submitted to an outlier analysis (Van Selst & Jolicoeur, 1994) that eliminated 2.44% of observations from subsequent analyses. Mean RTs were computed from the remaining observations. These mean RTs and corresponding error rates were submitted to two-tailed paired samples *t* tests. Naming times (see Table 1) were faster for repeated (577 ms) than for not-repeated (629 ms) items, $t(23) = 6.45$, $p < .001$, $d = 1.87$, and there was no difference in error rates for repeated (.001) and not-repeated (.002) items, $p > .10$.

Table 1

Mean response times (ms) and error rates for the study phase

Experiment	Repeated	Not-repeated	RT Difference (Rep – Not-rep)
1 – Repetition	577 (.001)	629 (.002)	-52***
2	609 (.008)	638 (.011)	-29***
3			
Lag-0	569 (.002)	594 (.006)	-25***
Lag-1	616 (.006)	640 (.008)	-24***
4			
Lag-0	420 (.001)	452 (.008)	-32**
Lag-1	464 (.001)	477 (.006)	-13

Note: Table displays response times with error rates in parentheses. Asterisks next to the RT difference score indicate a significant difference in RTs between the repeated and not-repeated items.
** *p* < .01; *** *p* < .001)

Test phase. For this and all following test phase analyses, items responded to incorrectly during the study phase were excluded from the test phase analysis. As a single red word was shown at test, a false alarm rate could not be computed separately for repeated and not-repeated items. As such, hit rates were used as a measure of sensitivity when comparing repeated and not-repeated items. The mean proportions of “old” responses are displayed in Figure 2.

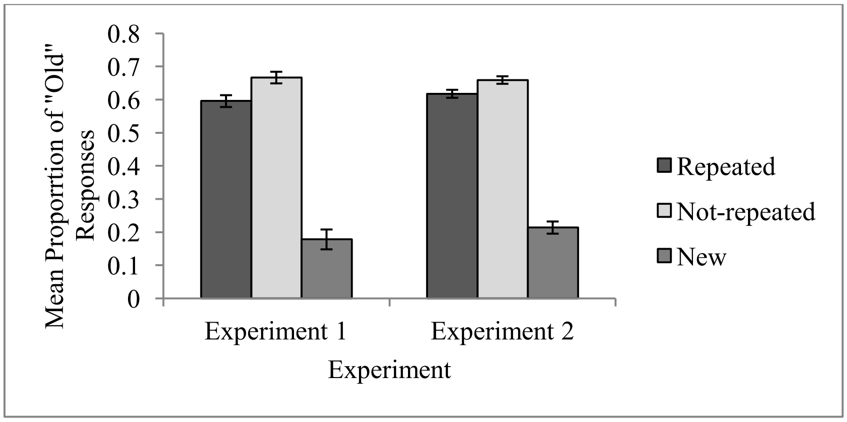

Figure 2: Mean proportions of “old” responses for Experiments 1 and 2. Error bars here and on all other graphs represent within-subject error corrected for between subjects variability (Morey, 2008).

A preliminary analysis was conducted that compared hits to false alarm rates, collapsed across repeated and not-repeated old items. This analysis was conducted simply to confirm that participants recognized items at above chance levels. A two-tailed paired sample *t* test confirmed that participants were able to distinguish between old (.631) and new items (.178), $t(23) = 12.3$, $p < .001$, $d = 3.56$. More important, a comparison of the hit rates revealed superior recognition of not-repeated (.667) than repeated items (.596), $t(23) = 4.62$, $p < .001$, $d = 1.33$.

The proportions of "remember" (R) and "know" (K) responses from the test phase were used to estimate the separate contributions of recollection and familiarity to recognition memory performance (Yonelinas, 2002; Yonelinas & Jacoby, 1995). The contribution of recollection is based on R, whereas the contribution of familiarity is based on K given an R response was not made (1-R). These estimates were obtained separately for hit rates for each item type.

The estimates of recollection and familiarity were analyzed separately using two-tailed paired sample *t* tests. The analysis of recollection estimates revealed no difference between repeated (.310) and not-repeated items (.330), $t < 1$. However, the comparison of familiarity estimates revealed higher familiarity for not-repeated (.512) than repeated items (.426), $t(23) = 4.39$, $p < .001$, $d = 1.27$.

Discussion

The results of Experiment 1 were counter to intuition: Recognition memory was more sensitive for not-repeated than repeated items, in this case driven by differences in familiarity rather than recollection (but see Collins, Rosner, & Milliken, 2018). This result is noteworthy in that it demonstrates better memory performance for an item presented once than for an item presented twice in succession at study. To our knowledge, no prior published study has reported such a result.

Experiment 2

The results of Experiment 1 are consistent with the desirable difficulty principle: the more difficult to name words at study (i.e., the not-repeated items) were recognized with greater sensitivity than the easier to name words (i.e., the repeated items). As noted in the introduction, this effect was discovered in the context of a research program focused on the link between adaptations in cognitive control and recognition memory. In particular, we were interested in the possibility that processing difficulty (or disfluency) might lead to an upregulation of attention, which would in turn improve memory encoding and recognition memory performance. The results of Experiment 1 are consistent with this idea, but of course other theoretical accounts are also tenable.

The purpose of Experiment 2 was to address an alternative account of this result. In particular, at the time of test, participants were instructed to identify words as “old” if they remembered reading them at the time of study. In the repetition group, a participant may have experienced the word “BRICK” as a repeated item at study, meaning that this word would have been presented first in green and then again in red. At test, the word “BRICK” would then appear in red, at which time the participant might remember reading it as a red target, seeing it as a green prime, or both. If the participant was of the mind that words presented in green at study disqualify an item as also having been presented in red, then remembering a green prime would have led participants to classify

"BRICK" as a new item. If this were the case, then performance would be systematically negatively affected for repeated relative to not-repeated trials, but not because of the influence of repetition on memory encoding. Instead, this result would be related to an idiosyncratic source memory judgment occurring at test.

Experiment 2 was conducted to rule out this particular source memory account. In Experiment 2, participants again completed an incidental study phase that involved naming the second of two words presented in sequence. However, rather than naming a red target word that followed a green prime word during the study phase, all study phase words (and all test phase words) were presented in white. As such, color could no longer be used as a basis for differentiating primes and targets at the time of test. If source memory for words of different color led to the results of Experiment 1, then the memory benefit for not-repeated words should not occur here. On the other hand, if processing difficulty differences related to stimulus repetition led to the results of Experiment 1, then the memory benefit for not-repeated words may be observed here as well.

Method

Participants. Forty-eight participants (37 female, mean age = 19, *SD* = 1.33) were recruited from the McMaster University student pool in exchange for course credit or $10. All participants spoke English fluently and had normal or corrected-to-normal vision.

Apparatus, stimuli, procedure, and design. The apparatus, stimuli, procedure, and design were identical to those for Experiment 1 with the following exceptions. All words were presented in white, all words were presented at the center of the screen (rather than offset slightly as in Experiment 1), and the spaces between the letters in each word that allowed interleaving in the congruency group of in Experiment 1 (see supplementary materials) were removed. During the study phase, participants were instructed to read the second of the two white words.

Results

Study phase. Correct RTs were submitted to the same outlier analysis used in Experiment 1, which removed 2.96% of RTs from further analysis (Van Selst & Jolicoeur, 1994). Mean RTs were computed from the remaining observations, and these mean RTs and error rates were analyzed using two-tailed paired sample *t* tests (see Table 1). RTs were faster for repeated (609 ms) than for not-repeated words (638 ms), $t(47) = 3.72$, $p < .001$, $d = 0.759$. No significant difference was observed for error rates, $p > .10$, but the error rates for the not-repeated words (.011) were numerically higher than for repeated words (.008), indicating no speed–accuracy trade-off.

Test phase. Two-tailed paired sample *t* tests were used to analyze data from the test phase. The overall hit rate, collapsed across the two repetition conditions, was first compared to the false alarm rate. This analysis revealed a higher proportion of “old” responses to old items (.638) than to new items (.214), $t(47) = 18.9$, $p < .001$, $d = 3.85$, indicating simply that participants were able to distinguish between old and new items. More important, an analysis of the hit rates revealed better recognition memory performance for not-repeated (.659) than repeated items (.618), $t(47) = 3.49$, $p = .001$, $d = 0.712$ (see Figure 2).

Estimates of recollection and familiarity were again compared for repeated and not-repeated items using two-tailed paired sample *t* tests. Reminiscent of the results of Experiment 1, familiarity estimates were higher for not-repeated (.481) than repeated items (.434), $t(47) = 3.78$, $p < .001$, $d = 0.771$. Again, there was no difference in recollection between repeated (.332) and not-repeated (.348) items, $t(47) = 1.18$, $p = .243$, $d = 0.241$.

Discussion

The results of Experiment 2 indicate that recognition sensitivity was higher for not-repeated than repeated items, as observed in Experiment 1. This result is not consistent with a source memory account by which old repeated items are incorrectly classified as “new” simply because they are remembered as being green rather than red at the time of study. Rather, it seems that the increased fluency experienced when responding to repeated targets may result in poor encoding of those items. As a result, these items are not as well-remembered as the not-repeated targets, as reflected by performance on a recognition memory task. Of course, it remains possible that participants can recall a word as having been either a prime or a target at study when making recognition decisions at the time of test. If so, then these source judgments could still support a source memory account of the superior recognition of not-repeated items. As such, we cannot rule out all possible source memory accounts of the present findings, but can rule out what we think of as an idiosyncratic source memory account related to the colours in which items are presented at study.

Experiment 3

The results of Experiments 1 and 2 demonstrate superior recognition for words presented once at study (not-repeated items) than words presented twice at study (repeated items). This result is surprising in that one might assume multiple opportunities to encode an item would benefit rather than hurt retention. However, in the current procedure, repeated words may not have been given the opportunity to be encoded twice. That is, although there were two physical presentations of the same word on repeated study trials, the short time interval between repetitions may have led the two repetitions to be encoded as a single event (Hommel, 1998, 2004; Kahneman, Treisman, & Gibbs, 1992).

The goal of Experiment 3 was to examine whether the memory benefit for not-repeated words observed in Experiments 1 and 2 is related, at least in part, to the repeated study items being perceived as a single event. To address this issue, participants were separated into two groups: a lag-0 group and a lag-1 group. All participants were presented with three words (two primes and a target) in succession on every study trial (see Figure 3). For not-repeated items, these three words were all different. For repeated items, two of the three words had the same identity. Repeated items for the lag-0 group contained an identical word in serial positions two and three in each three-word study trial. Given that repetition in this condition was not spaced (i.e., there were no items intervening between the repeated items), the prediction was that performance should be similar to that in Experiments 1 and 2, with better recognition for not-repeated than repeated items. In contrast, repeated items for the lag-1 group contained the same word in positions one and three in each three-word study trial. A different word appeared in serial position two, which in effect spaced the repetitions by a single item. If the results of Experiments 1 and 2 were a by-product of repeated words being encoded as a single event (e.g., Hommel, 2004), then the results from the lag-1 group should differ from those of Experiments 1 and 2. Indeed, if the intervening prime in the lag-1 condition fosters two separate encodings of the same word, then memory performance might well be better for repeated than for not-repeated trials.

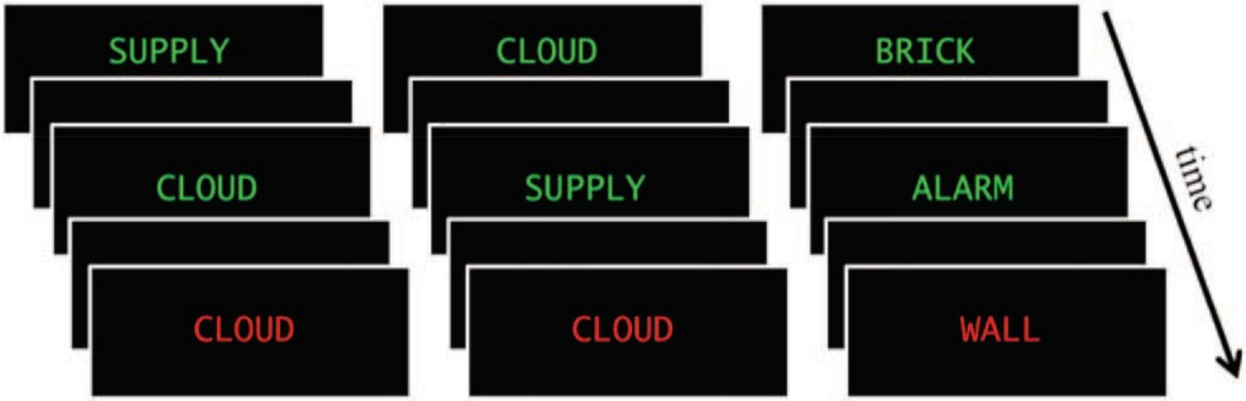

Figure 3. Examples of the trial sequence in Experiments 3 and 4. The leftmost sequence is an example of a repeated item for the lag-0 group; the middle sequence is an example of a repeated item for the lag-1 group. The rightmost sequence is an example of a not-repeated item, which was the same across groups. See the online article for the color version of this figure.

Method

Participants. Forty-eight participants (29 female, mean age = 20, *SD* = 1.84) were recruited from the McMaster University student pool in exchange for course credit or $10, or by undergraduate students as part of a third-year Memory and Cognition Laboratory course. These participants were randomly assigned to either the lag-0 group or the lag-1 group, with 24 participants in each group. All participants spoke English fluently and had normal or corrected-to-normal vision.

Apparatus and stimuli. The apparatus and stimuli were identical to Experiment 1 with the following exceptions. All words were presented at the center of the screen and the spaces between the letters in each word were removed. In addition, four- to six-letter high frequency nouns were used to construct the word lists and can be viewed in Appendix B (Thorndike & Lorge, 1944).

Procedure. The procedure was identical to Experiment 1 with the following exceptions. Every trial during the study phase consisted of the presentation of three words: two green primes and one red target. Each prime was presented for 500 ms, the target was presented for 1,000 ms, and there was an ISI of 250 ms between each word. The participants were instructed to read the red target word only. For the lag-1 group, on repeated trials the first of the two prime words and the target word were the same, whereas the second prime word differed from the other two words; that is, these repeated trials involved repetition of a word with an intervening different word. On the other hand, for the lag-0 group, the second prime word and the target word were the same, whereas the first word differed from the other two words. In this case, repeated trials involved repetition of a word with no intervening different word (as in Experiments 1 and 2). Not-repeated trials consisted of three different words, and were identical across the lag-0 and lag-1 groups (see Figure 3). Finally, the test phase required an old/new judgment as in prior experiments, but did not include a remember/know judgment.

Design. The design of Experiment 3 was identical to Experiment 1, with the exception that eight word lists of 60 words were used to construct the stimuli. An additional two word lists were added to create a filler prime word list. The assignment of each list to the repeated word, not-repeated target, not-repeated prime, filler prime, and new roles was counterbalanced across participants.

Results

Study phase. Correct RTs were submitted to an outlier analysis, which excluded 2.87% of observations from further analyses (Van Selst & Jolicoeur, 1994). Mean RTs were computed from the remaining observations, and these mean RTs and corresponding error rates were submitted to a 2 X 2 mixed-factor analysis of variance (ANOVA), with group (lag-0/lag-1) as a between-subjects factor, and repetition (repeated/not-repeated) as a within-subject factor (see Table 1). The analysis of RTs revealed a significant main effect of repetition, $F(1, 46) = 30.9$, $p = .001$, $\eta_p^2 = .402$, with faster RTs for repeated (592 ms) than not-repeated words (617 ms). Neither the main effect of group nor the interaction between group and repetition were significant, $p > .10$, indicating no difference in repetition effect between groups. The analysis of error rates produced similar results, revealing a repetition effect that approached significance, $F(1, 46) = 3.75$, $p = 059$, $\eta_p^2 = .075$, with higher error rates for not-repeated (.007) than repeated words (.004). Once again, all other effects were not significant, $p > .10$.

Test phase. To ensure that participants were able to distinguish between old and new items, the mean proportions of “old” responses, collapsed across the two repetition conditions, were submitted to a 2 X 2 mixed-factor ANOVA, with group (lag-0/lag-1) as a between-subjects factor and item type (old/new) as a within-subject factor. A main effect of item type was observed, $F(1, 46) = 229$, $p < .001$, $\eta_p^2 = .833$, with a higher proportion of "old" responses for old (.541) than new items (.208). No other effects in this analysis were significant, all *ps* > .10. Memory sensitivity was analysed by submitting hit rates for repeated and not-repeated items to a 2 X 2 mixed-factor ANOVA, with group (lag-0/lag-1) once again as a between-subjects factor, and repetition (repeated/not-repeated) as a within-subject factor (see Figure 4). This analysis revealed a significant main effect of repetition, $F(1, 46) = 13.3$, $p < .001$, $\eta_p^2 = .224$, with a higher hit rate for not-repeated (.568) than repeated items (.514). Importantly, neither the main effect of group, nor the interaction between group and repetition were significant, all *ps* > .10.

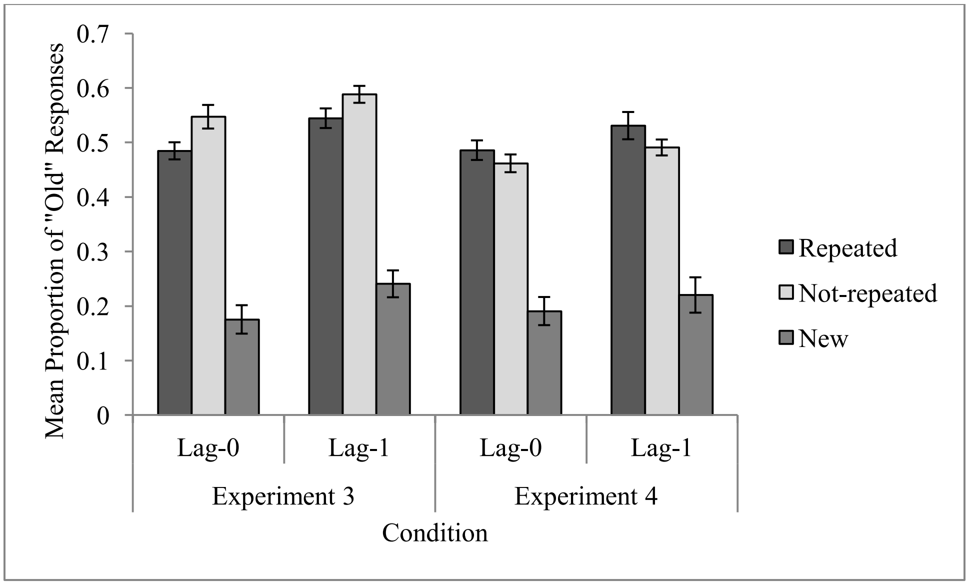

Figure 4: Mean proportions of "old" responses for Experiments 3 and 4.

Although there was no significant interaction between group and repetition, our a priori hypothesis that there could well be a difference between the groups led us to examine the effect of repetition separately for the two groups. Hit rates for repeated and not-repeated items were submitted to two-tailed paired sample *t* tests for the lag-0 and lag-1 groups separately. For the lag-0 group, a significant effect was found $t(23) = 2.79$, $p = .010$, $d = 0.807$, with a higher hit rate for not-repeated (.547) than repeated items (.485). For the lag-1 group the same result was observed; the hit rate was higher for not-repeated (.588) than for repeated items (.544), $t(23) = 2.34$, $p = .028$, $d = 0.676$.

Discussion

The results of this experiment were very clear. Recognition sensitivity was higher for not-repeated than for repeated items, and this effect on recognition sensitivity did not depend significantly on whether there was an intervening different word between the two repeated words.

Experiment 4

The results of Experiment 3 demonstrated that the superior recognition for not-repeated over repeated items observed with the present procedure is robust to the presentation of an intervening different item between a repeated prime and target. This result prompted additional thought about factors that were unique to our method that might produce better memory for an item presented once than an item presented twice. We noted that studies of the spacing effect typically involve attentive encoding of the same event twice, either unspaced or spaced (e.g., Baddeley & Longman, 1978; R. A. Bjork & Allen, 1970; Kahana & Howard, 2005), whereas our procedure required participants to attentively encode only the second of two repeated events. To address this issue in Experiment 4, participants again saw two green prime words and a red target word, and were placed either in the lag-0 or lag-1 group. Importantly however, participants were asked to read aloud all items in the study phase. If the results of previous experiments were in part due to the primes being ignored rather than attended to, then Experiment 4 should produce either a null effect or the opposite result, namely, better memory for repeated than not-repeated words.

Methods

Participants. Forty-eight participants (39 female, mean age = 20, *SD* = 2.54) were recruited from the McMaster University student pool in exchange for course credit or $10. These participants were randomly assigned to the lag-0 or lag-1 group, with 24 participants in each group. All participants spoke English fluently and had normal or corrected-to-normal vision.

Apparatus, stimuli, procedure, and design. The apparatus, stimuli, procedure, and design were identical to Experiment 3 with the following exceptions. Participants were instructed to read all three words on every trial as they appeared, though RTs were only recorded for the final target words. In addition, the ISI between words was increased from 250 ms to 500 ms to ensure that participants had enough time to read each word before the next one appeared.

Results

Study phase. Correct RTs were submitted to the same outlier analysis as in previous experiments which eliminated 2.49% of observations from further analyses

(Van Selst & Jolicoeur, 1994). Mean RTs were computed from the remaining observations and these mean RTs and corresponding error rates were submitted to the same 2 X 2 mixed-factor ANOVA as in Experiment 3, with group (lag-0/lag-1) as a between-subjects factor and repetition (repeated/not-repeated) as a within-subject factor (see Table 1). The analysis of RTs revealed a significant main effect of repetition, $F(1, 46) = 8.94$, $p = .005$, $\eta_p^2 = .163$, with faster RTs for repeated (442 ms) than not-repeated words (464 ms). A significant effect of repetition was also observed in the analysis of error rates, $F(1, 46) = 10.3$, $p = .002$, $\eta_p^2 = .183$, with higher error rates for not-repeated (.007) than repeated words (.001). All other effects in both analyses were not significant, all *ps* > .10.

Test phase. As in Experiment 3, the mean proportions of "old" responses to old items collapsed across the two repetition conditions, and to new items, were submitted to a 2 X 2 mixed-factor ANOVA, with group (lag-0/lag-1) as a between-subjects factor and item type (old/new) as a within-subject factor. The analysis revealed a main effect of item type, $F(1,46) = 127$, $p < .001$, $\eta_p^2 = .735$, with a higher proportion of "old" responses to old (.492) than new items (.205); no other effects were significant, all *ps* > .10. Next, an analysis of hit rates was conducted using a 2 X 2 mixed-factor ANOVA, with group (lag-0/lag-1) as a between-subjects factor and repetition (repeated/not-repeated) as a within-subject factor (see Figure 4). This analysis revealed a main effect of repetition, $F(1, 46) = 5.45$, $p = .024$, $\eta_p^2 = .106$, with a lower hit rate for not-repeated (.476) than repeated items (.508). Note that this result is opposite to those reported in all of the other experiments reported here. No other effects were significant in the analysis, all *ps* > .10.

Although the interaction between group and repetition was not significant, the effect of repetition was analyzed separately for the two lag groups for the sake of consistency with the analyses reported for Experiment 3. For both the lag-0 and lag-1 groups, hit rates for repeated and not-repeated items were analyzed using two-tailed paired sample *t* tests. For the lag-1 group, the effect of repetition approached significance, $t(23) = 1.96$, $p = .062$, $d = 0.567$, with a higher hit rate for repeated (.531) than not-repeated items (.491). For the lag-0 group, the hit rate difference between repeated (.486) and not-repeated (.462) items was not significant, $t(23) = 1.31$, $p = .204$, $d = 0.378$.

Discussion

The results of Experiment 4 were opposite to all others reported in the current study: recognition memory sensitivity was greater for repeated than not-repeated items. This result points to the idea that superior recognition for not-repeated over repeated items may hinge on inattention to primes (see also Collins et al., 2018). Also of note is that the benefit of repetition in the present experiment was not a particularly strong one; hit rates were only about 3% higher for repeated than not-repeated words. Together with the results of previous experiments, this result is consistent with an opponent process view; that is, the results in this experiment may reflect the joint benefit of repetition associated with attending to primes and a cost of repetition that is unrelated to attending to primes.

General Discussion

The goal of the current study was to examine the effect of immediate repetition on recognition memory. In Experiment 1, a green word prime followed by a named red word target produced poorer recognition for repeated than not-repeated words. To our knowledge, this specific finding has not previously been reported, yet it is a result we observed reliably across several experiments. In Experiment 2, the same result occurred with identically colored white prime and target words, demonstrating that the original effect was not an artifact of source memory for word color at study. Another possible explanation for these counterintuitive findings was that repeated words were not given a proper chance to be encoded as two separate events. This account was ruled out in Experiment 3, in which better memory was found for not-repeated than repeated words even when there was an intervening different word between the two repeated words. Finally, in Experiment 4, when both the prime and target words were named aloud, better memory was found for repeated than not-repeated items, demonstrating that immediate repetition effects on recognition depend on how primes are encoded. An important objective now and going forward will be to consider the broader implications of these findings for our understanding of the influence of stimulus repetition on perception, learning, and memory.

Our preferred view of the key new result reported here makes reference to the role of attentional orienting in event encoding; we orient preferentially to new events rather than old events (Klein, 2000; Posner & Cohen, 1984; Yantis & Jonides, 1984)[3].

---

[3] Although the inhibition of return effect is most often measured with respect to shifts of attention in space, a similar mechanism may be responsible for slow responses to repetitions of nonspatial stimulus dimensions, such as colour (Klein, 2000; Law, Pratt, & Abrams, 1995; Spadaro et al., 2012). Thus, we view it as possible that inhibition of return mechanisms are operating here, where effects of repetition for stimulus identity rather than stimulus location are the focus. Furthermore, it should be noted that inhibition of return is observed with long ISIs, whereas the current experiment uses short ISIs (250 and 500 ms). However, it is possible that the mechanism that responds to novelty (leading to inhibition of return effects) is still operative with short ISIs. That is, even at short ISIs a response to novelty may still occur, though this

Preferential shifts of attention to novel events may lead to better event encoding of target words that mismatch the preceding primes than for target stimuli that match preceding primes. This core idea offers a possible foundation for explaining the results reported here, but of course there are other possibilities as well. In the remainder of the article we discuss alternative interpretations of the results, and point to future research paths that appear promising.

The Role of Prime Encoding

One result worth discussing is that immediate repetition effects on recognition memory depended on how primes in the study phase were encoded. There are at least two distinct ways of thinking about this result. By one view, the dependence of repetition effects on prime encoding focuses on whether the prime encoding demands match or mismatch those of the following target, whereas the other view focuses on the extent to which attention is actively allocated to the primes. These two ideas are unpacked in more detail below.

A transfer appropriate processing account. One account of the dependence of immediate repetition effects on prime encoding focuses on the transfer appropriate processing principle (Morris, Bransford, & Franks, 1977). When applied to prime-target procedures like that used here (Neill & Mathis, 1998; Wood & Milliken, 1998), this principle states that transfer of processing from prime to target depends on the specific correspondence between how prime and target are encoded. Thus, when participants were

---

response may not be observed in RTs at short ISIs because of an opposing mechanism producing facilitation for repetitions (Klein, 2000).

asked to ignore a prime and then name a following target, the processing requirements from prime to target mismatched. As a consequence, on repeated prime-target study trials, the match in word identity together with the mismatch in processing requirement (i.e., ignore and then name) would be expected to produce inefficient transfer of processing from prime to target (see also Hommel, 1998, 2004). It follows that this "transfer inappropriate processing" on repeated study phase trials may have disrupted memory encoding for targets and produced poor recognition at test. In line with this view, when participants named both the prime and target in Experiment 4, transfer from prime to target would instead be "appropriate," which accounts for why recognition sensitivity was higher for repeated than not-repeated trials.

A challenge for this account of the present results is that naming times in the study phase were faster for repeated than for not-repeated trials regardless of how the prime was encoded. For the transfer appropriate processing account to hold, one would have to argue that test phase recognition is subject to transfer appropriate processing during the study phase, whereas study phase naming times are driven by some other process. Another notable implication of the transfer appropriate processing account is that poor recognition is attributed to processing difficulties at the time of target encoding, which runs counter to the desirable difficulty principle. Recall that the desirable difficulty principle instead attributes good rather than poor memory performance to processing difficulties at the time of target encoding.

An opponent process account. An alternative account of the prime encoding effects observed in the present study focuses on the degree to which participants attend to the identity or meaning of the primes. When participants attend to primes, memory for the prime itself might support recognition of repeated items as old at the time of test. For example, naming the prime word CLOUD prior to naming an identical target word CLOUD could result in participants making an "old" response to CLOUD at test either because they remember the target episode or because they remember the prime episode. From this perspective, attending to primes introduces an additional source of evidence that favors a correct "old" response at test for repeated items from the study phase and helps explain the results observed in Experiment 4 of this article. In contrast, ignoring primes makes it more likely that "old" responses for repeated items are driven only by memory of the target episode (i.e., ignoring a word at study is an ineffective way to remember it at test). If we also assume that memory for target episodes benefits from more robust attention orienting to not-repeated than repeated targets during the study phase, then this would account for the recognition results observed in Experiments 1–3 of this article.

Furthermore, the opponent process account is consistent with the desirable difficulty principle (R. A. Bjork, 1994), which may help to explain the results of Experiments 1–3, in which recognition sensitivity was higher for not-repeated than repeated trials. According to the desirable difficulty principle, greater encoding difficulty in the study phase resulted in better recognition memory in the test phase, an effect consistent with other reported perceptual desirable difficulties (e.g., Diemand-Yauman et al., 2011; Hirshman & Mulligan, 1991; Nairne, 1988; Rosner, Davis, et al., 2015). On the other hand, when primes were actively attended to (i.e., by reading them, as in

Experiment 4), the results were not in line with the desirable difficulty principle. These results are easily accommodated by the opponent process account, which suggests that the desirable difficulty principle was still operative, but that recognition performance was also controlled by a second process related to attention to the primes that pushed performance in the opposite direction.

Evaluating the role of prime encoding. Collins et al. (2018) recently conducted a series of experiments to address prime encoding issues with a procedure similar to the one used in the present study. In some conditions, participants viewed the prime passively (either by being told explicitly to ignore the prime, or by dividing attention during prime presentation); in others, the prime was actively encoded (either by reading or semantically encoding the prime). In all conditions, participants were asked simply to read the red target words, which were then later seen in a recognition memory task. In the passive encoding conditions, notrepeated words were better remembered than repeated words, as in Experiments 1–3 reported here. In the active conditions, particularly so with semantic encoding, the opposite effect was found, with better memory for repeated than not-repeated words. Overall, together with the results reported here, it seems that the memory benefit for not-repeated items hinges on poor encoding of, and therefore poor subsequent memory for, the prime episode. This conclusion fits well with the opponent process account proposed here. Furthermore, Collins et al. reported results that were inconsistent with the transfer appropriate processing account, as substantially better memory for repeated than not-repeated items was observed when processing of the prime and target differed (i.e., semantically encode the prime and read the target).

There are a variety of additional ideas related to prime encoding that remain to be tested. For example, in the present study we contrasted performance in a repeated and not-repeated condition and have proposed that performance may benefit both from attentional orienting to the encoding of not-repeated targets and from actively attending to repeated primes. This opponent process proposal may benefit from further tests that include a neutral condition, perhaps one in which participants name a single red word that follows a meaningless string of prime characters. This method may help to distinguish between benefits and costs of repetition on recognition performance, as well as how these benefits and costs vary as a function of changes in prime encoding.

Another issue that has yet to be examined is whether briefly presented and masked primes would produce a similar pattern to ignored primes presented for a longer duration, as in the present study. The opponent process account rests on the idea that attention orients more robustly to novel objects than familiar objects, which improves memory encoding for not-repeated relative to repeated targets. If this preferential orienting to novelty is driven by a comparison of perceptual representations of the target and preceding prime, then we would predict that not-repeated items would be better recognized than repeated items even if the primes were presented too quickly to be consciously perceived. However, the effect observed in the current set of experiments may require higher-level representations of primes, which may not be accessible in a study in which primes are quickly presented and masked (Henson & Gagnepain, 2010). This is an issue well worth future study.

Issues of Spacing

In Experiments 3 and 4 of the current study, spacing between repetitions was manipulated to ensure that repeated items were experienced as two separate instances. It could be argued, however, that the results of Experiments 3 and 4 constitute a failure to find the well-known spacing effect, in which spaced repetitions are better remembered than massed repetitions (Kahana & Howard, 2005). According to the spacing effect principle (Baddeley & Longman, 1978; R. A. Bjork & Allen, 1970), at least insofar as a single intervening item constitutes an effective manipulation of spacing, the repeated words in the lag-1 group should have been better remembered than the repeated words in the lag-0 group. Yet, in both experiments, there was no effect of this spacing manipulation.

This lack of a spacing effect may seem troubling, although there are a few reasonable explanations for why these experiments did not show such an effect. First, most spacing effect studies (e.g., Cuddy & Jacoby, 1982; Hintzman, 1974; Jacoby, 1978; Kahana & Howard, 2005; Underwood, 1969) tend to look at larger intervals between repetitions (i.e., more than one intervening item). As such, it may be that a spacing effect was not found here simply because the interval between repetitions was too short. Second, some studies have shown that even with intervening items between repetitions, engaging in the same task (e.g., reading) for those repetitions results in little or no memory benefit (Cuddy & Jacoby, 1982; Jacoby, 1978). Therefore, it is possible that even if there were larger intervals between repetitions in the current study, the memory benefit afforded by increased spacing may be minimal. A systematic study of this issue that compares immediate repetition effects as measured here, to repetition effects across larger spacing intervals, is certainly warranted.

It should also be noted that few spacing effect studies have compared memory for single presentations of words to massed repeated presentations of words, as was the case here[4]. Those studies that have addressed this issue have provided mixed results. Some experiments have produced better memory for massed (i.e., repeated) than single (i.e., not-repeated) presentation items, whereas experiments in the same set of studies have produced equivalent memory performance for the two conditions (Braun & Rubin, 1998; Cuddy & Jacoby, 1982; Underwood, 1969). In addition, these studies did not explicitly examine immediate repetition effects (e.g., repetition over an interval of less than a second), making it unclear how stimulus repetition effects in those studies compare to those in the present study. Although to this point we have been unable to set our result (superior memory for not-repeated over repeated items) precisely in the context of the broader literature on spacing effects, our result does point to an important gap in the literature that requires further study.

---

[4] Studies of the spacing effect compare memory performance for repetitions of study items spaced closely (massed) or spaced further apart, whereas the effect we report here compares memory for items presented just once to that for repetitions of items spaced closely together at study. As such, there is something fundamentally different between the spacing effect and the effect we report here. However, studies examining immediate serial recall for simple stimuli (such as numbers or letters) have compared memory for an item presented once with memory for repeated items. Specifically, spacing of repetitions in immediate serial recall makes memory worse rather than better, producing the opposite of the spacing effect. This finding is known as the Ranschburg effect (Crowder & Melton, 1965; Jahnke, 1969b; Lee, 1976). In these studies, participants see a string of eight to 10 items; control strings contain all different items, whereas repetition strings contain one repeated item. Immediate serial recall for the repeated items is worse than for comparable control items when the repetitions are spaced but not when they are unspaced. Although this effect demonstrates a memory cost for repeated items, it appears to be different from the effect examined in the current study, as it is presumed to be due to processes at the time of recall such as output interference (Jahnke, 1969a) or guessing strategies (Greene, 1991) rather than to attention orienting to novel items and an up-regulation of attention in response to processing difficulty. (See also repetition blindness for a related effect; Kanwisher, 1987).

The Repetition Effect and the Word Frequency Effect

Recognition sensitivity is typically higher for low frequency words than for high frequency words (Gregg, 1976; see also Gillund & Shiffrin, 1984; Glanzer & Adams, 1985; Higham, Perfect, & Bruno, 2009; Joordens & Hockley, 2000; Kinoshita, 1995; Kinsbourne & George, 1974). This well-studied regularity of recognition may be relevant to the results reported here. Although all of the items used in our study were high-frequency words, and therefore the repetition effects observed here are not procedurally equivalent to word frequency effects, it is worth discussing possible similarities in the underlying mechanisms that may produce the repetition and word frequency effects.

An interesting possibility is that study phase naming of repeated targets in our experiments is more efficient than for not-repeated targets because of the high accessibility of a compatible memory representation from the immediately preceding prime. Similarly, high frequency words may be named more efficiently than low frequency words because of more readily accessible memory representations for words that occur more frequently in natural language. In other words, the same fundamental constraint of memory search for an appropriate target representation could underlie both the repetition and word frequency effects. Expanding this idea a step further, when a memory representation is not immediately accessible (i.e., for low frequency or not-repeated targets) additional processing may be required that involves refining the target representation, so that a search of memory is successful despite the memory representation being of relatively low accessibility (see Hintzman, 1988; Jamieson, Holmes, & Mewhort, 2010). As such, processes related to refining the target representation may result in attention orienting that is more robust for not-repeated than for repeated targets. Future research that examines the immediate repetition effects reported here within this general framework is certainly needed.

Another interesting focus for future research is whether the immediate repetition effects observed here are specific to recognition or if they generalize to other memory tasks. Although low frequency words are better recognized than high frequency words, the opposite is typically observed in free recall tests (Gregg, 1976; see McDaniel & Bugg, 2008 for a review). If the repetition effect reported here is related to the word frequency effect, then recall performance may be better for repeated than not-repeated words. This result would offer important converging evidence in favor of the view that repetition and word frequency effects have the same underlying cause.

Time-On-Task

One other account for the present findings is the possibility that the more time spent processing an item, the better it is remembered. In other words, better recognition memory performance for not-repeated than repeated items in the current study may simply be a direct result of longer processing times for not-repeated than repeated targets. That is, the more time spent processing an item directly results in that item being better remembered. It is important to note that this explanation is inconsistent with the results of Experiment 4, in which not-repeated targets were processed more slowly than repeated targets, and yet recognition sensitivity was greater for repeated than not-repeated items. Nevertheless, it is still possible that time-on-task may explain the results of Experiments

1–3, in which the not-repeated targets required more processing than repeated targets, and were also better remembered at test.

To address this concern, we compared RTs for remembered and forgotten items separately for repeated and not repeated targets. If time-on-task is responsible for the results of the present study, then RTs should be slower for remembered than forgotten items. One-tailed *t* tests indicated that this was not the case for both repeated and not-repeated items, both *t*'s < 1. More critically, we also compared RTs for remembered repeated targets and forgotten not-repeated targets. Again, if time-on-task is a key issue in this study, then RTs should be slower for remembered repeated targets than for forgotten not-repeated targets. Further analyses revealed that this was not the case, as RTs for remembered repeated targets (595 ms) were faster than for forgotten not-repeated targets (620 ms), $t(118) = 4.76$, $p < .001$. These analyses demonstrate that more time spent processing an item does not directly lead to that item being well-encoded.

Conclusions

The present study describes a novel finding in which an immediate repetition method at study revealed better recognition for not-repeated than repeated items at test. This result was observed across multiple experiments and has been replicated in our laboratory numerous times subsequently (see also Collins et al., 2018). We propose that this effect points to a fundamental issue at the interface between attention and memory encoding that merits additional study: Is this effect related to the well-known spacing and word frequency effects in the memory literature (Baddeley & Longman, 1978; R. A. Bjork & Allen, 1970; Joordens & Hockley, 2000)? Is this effect related to attention orienting phenomena that implicate more robust orienting to novel than familiar events (Posner & Cohen, 1984; Yantis & Jonides, 1984)? These issues converge on a central theme that perceptual encoding is adapted flexibly in response to how well-predicted external events are by memory representations of prior experiences (Henson & Gagnepain, 2010). Critically, our findings indicate that encoding may potentially be negatively impacted by repetition and may be related to other literatures in which poor performance is observed in response to repetitions (Milliken, Joordens, Merikle, & Seiffert, 1998; Posner & Cohen, 1984; Spadaro & Milliken, 2013; Tipper, 1985; Yantis & Jonides, 1984).

Appendix A

Word lists: Experiments 1 and 2

Word list 1: CURVE, MONEY, TOWER, WHEEL, TABLE, CHAIR, GRATE, GROUP, STEEP, STERN, DAILY, TOTAL, WATCH, PILOT, WATER, SHORT, NOVEL, MONTH, SHEER, CABLE, GUIDE, CATCH, BRIDE, CLEAN, TITLE, SPRAY, SOLID, MAJOR, CLIMB, THROW, COAST, OLIVE, RIVAL, TRUTH, IDEAL, GLOVE, STAMP, JEWEL, ADULT, PURSE, GUARD, SWIFT, WORST, CHARM, UNCLE, SIGHT, FLASH, STORY, JUICE, TRADE, PITCH, SAUCE, FANCY, SHOCK, STRIP, BLIND, BROOK, START, GLARE, OPERA

Word list 2: TOUGH, CLERK, PEACH, PHONE, KNOCK, OFFER, GLEAM, WOUND, SPEED, PUPIL, LIGHT, VALUE, EVENT, CREEK, RANCH, CRASH, TRACE, MOTOR, SHAPE, SMART, EXTRA, SHIRT, VOICE, ANGLE, STAKE, RADIO, MATCH, TODAY, BRICK, SLOPE, BREAD, SWEAR, PARTY, BREAK, CHILL, TRAIN, UPPER, WORLD, GLASS, MOVIE, PROOF, NOBLE, STRAW, SCORE, BASIS, CLASS, EMPTY, STATE, AGENT, MAGIC, FRONT, FROWN, CABIN, BIRTH, FLOOR, CIGAR, PIANO, CHOKE, SLIDE, COUNT

Word list 3: TWIST, FEVER, TRICK, DELAY, TOAST, SPOON, DREAM, CHEST, STUFF, CRAWL, LUNCH, INNER, TASTE, BENCH, NURSE, CHAIN, NERVE, RANGE, ISSUE, CLOUD, CHASE, HONEY, HORSE, PLANE, OWNER, LIMIT, PRESS, ROUND, PAINT, PRIZE, ASIDE, CANDY, TREAT, BLANK, SHAME, STOOP, MOUTH, FLAME, ANKLE, BATHE, LAUGH, MUSIC, SCALE,

POUND, OCEAN, MIGHT, CLAIM, FAINT, YIELD, CHIEF, HEART, ONION, CHEER, BRAND, PLANK, SLEEP, STUDY, TENSE, GUESS, LEAST

Word list 4: STOLE, SHAKE, CHILD, ROUGH, CLIFF, ORDER, REBEL, DRESS, ELBOW, WAGON, SHARP, QUOTE, TIMER, TRAIL, CROWN, STAGE, JELLY, LINEN, DRAIN, STALK, HURRY, PORCH, RIGHT, STEEL, BRUSH, TRUNK, BUNCH, PENNY, BLOCK, SERVE, STONE, SHARE, BRAIN, EARTH, SCENE, FLOUR, CHEEK, GRASS, PLANT, JUDGE, CRUMB, PRINT, SPOIL, SPOKE, SHELL, OTHER, COURT, SHEET, PAUSE, SUGAR, STICK, BLOOM, TEETH, PRIDE, WHILE, SKIRT, BLAZE, GLORY, TRACK, SPELL

Word list 5: CORAL, FRAME, SHORE, GRAIN, STORE, BOAST, DOUBT, SATIN, REACH, SLICE, PEARL, GRASP, PRICE, ORGAN, TRUCK, STOCK, PASTE, CRACK, COVER, SWING, WHIRL, CLOCK, RIVER, SPACE, STEAL, LEVEL, DEPTH, STILL, ROUTE, SPLIT, SCARE, FENCE, KNIFE, ACTOR, POINT, THING, FLOAT, SALAD, GRIEF, SHINE, SMELL, QUIET, SHIFT, SCENT, LEMON, ELECT, FRUIT, GUEST, MIDST, FLUSH, PIECE, DRIVE, GRADE, SWEET, QUICK, NOISE, SMALL, CROSS, STAND, TROOP

Word list 6: VISIT, APPLE, STYLE, FIELD, BOUND, SWEAT, METAL, LEAVE, DRINK, WRIST, THUMB, MORAL, DANCE, STARE, GRANT, POISE, STOVE, GROAN, SOUND, HOUSE, SHOUT, DRIFT, SENSE, CLOTH, CROWD, LAYER, STORM, WASTE, SMILE, ROAST, SHRUG, PLATE, TRUST, BLUSH, CRUSH, COACH, HOTEL, PAPER, YOUTH, CHECK, SAINT, WRECK, SPORT,

EQUAL, SMOKE, STAFF, BURST, BOARD, LOCAL, STEAM, FORCE, ALARM, SHADE, NIGHT, WOMAN, MODEL, UNDER, WHEAT, BRIEF, TOUCH

Appendix B

Word lists: Experiments 3 and 4

Word list 1: JAZZ, CLOSE, WINDOW, FOREST, FATHER, MEANS, LEAST, TRADE, SIDE, SECRET, SENSE, NOVEL, SORT, VISIT, SWEET, EFFECT, CHANCE, DESIGN, SOFT, VOICE, PARTY, BRIDGE, BLOOD, DESK, FEEL, APRIL, FAMILY, STILL, DOCTOR, INCOME, WRONG, COLD, MODERN, ORDER, SOUTH, VOLUME, FINE, BLOCK, LACK, EXTENT, FRIDAY, RIGHT, FILM, IMPACT, CHECK, SOURCE, COUNTY, GROUND, SHARE, PAST, REACH, BEAUTY, PARK, THING, FLAT, MARCH, SIGNAL, SHORE, LEAD, SHAPE

Word list 2: FOOT, CLAIM, GOAL, PRICE, SEARCH, FIRM, LIVING, DINNER, HOLD, MOUTH, BODY, CHARGE, SHOT, FACE, ANIMAL, CARE, ISSUE, WORLD, CAREER, SONG, BATTLE, DEEP, MANY, PAPER, CHILD, VALLEY, ROSE, CORNER, TIME, APPEAL, BALL, ROLE, LIGHT, COMMON, SIZE, THEORY, IMAGE, SOVIET, JUDGE, HALF, LEAVE, YOUTH, FELL, PASS, IDEAL, VIEW, NATURE, FACT, TRAVEL, ASIDE, PART, LAST, WHILE, SHORT, PUBLIC, MIDDLE, SLEEP, DRIVE, MASS, STAFF

Word list 3: YEAR, GLASS, TOWN, DEGREE, THICK, DANCE, TITLE, MUST, FIELD, SECOND, SAFE, LINE, NOTICE, IDEA, PLAN, FIND, HEAD, MONTH, FEED, NONE, PAGE, FEAR, PHASE, UNDER, HOUSE, OFFICE, BOOK, LETTER, WANT, MOTHER, NORTH, SNOW, CIRCLE, ENERGY, FRIEND, PLANT, CLOSER, DRESS, FORM, NARROW, TYPE, FAST, BLUE, DOWN, MEET,

BEACH, ARMY, WATER, WOMEN, UPPER, MIGHT, FELLOW, CHIEF, MISS, WARM, STREET, CITY, HEART, HAIR, RISE

Word list 4: POINT, LOSS, SPIRIT, MOON, FALL, QUIET, START, STOP, SCHOOL, VALUE, DUTY, INSIDE, BUDGET, TODAY, CATTLE, PERSON, QUICK, ROUND, MIND, GOOD, TEST, CALL, SPACE, SHIP, ROOF, EVENT, SHOP, HELP, FILE, TONE, FEET, ISLAND, SOUND, CHAIR, MONEY, COURSE, MOMENT, POST, GAIN, HORSE, HOME, CENT, AMOUNT, SAVE, PIECE, MONDAY, SPREAD, JOIN, WELL, GROWTH, TOTAL, FAIR, DETAIL, DREAM, LONGER, CELL, COLUMN, DEMAND, KING, STAY

Word list 5: WIDE, INDEX, STOCK, STRESS, TALK, HAND, SUNDAY, TRIP, STUDY, FACTOR, VOTE, CRISIS, LOOK, EARTH, BELIEF, JUNE, STAND, SUMMER, JUNIOR, HOUR, ROCK, MATTER, MAKE, EDGE, COST, HAVE, BOARD, FRAME, LIKE, WEEK, GIRL, ROAD, SPEED, BOTTLE, HALL, CREDIT, BREAK, DRINK, TURN, RACE, BETTER, DOUBT, MAJOR, HEALTH, BIRTH, SUPPLY, WIND, SIGHT, MEMORY, LAND, SPRING, LIST, PRESS, SERIES, TRAIN, NOTE, SEASON, STEP, BASIS, RATE

Word list 6: NEWS, SCALE, SPEECH, MEMBER, DATE, SERVE, LEVEL, CAUSE, MOVE, JACK, AREA, HOPE, FRONT, SCENE, MAIN, SITE, LOCAL, TEETH, METHOD, LENGTH, NOSE, NOBODY, STYLE, MARKET, EFFORT, POLICY, WORK, DROVE, ROOM, WORD, JURY, FARM, CHANGE, BOAT, CENTER, TRUTH, HEAT, SIGN, RELIEF, EAST, MUSIC, STATE, WALK, PEOPLE,

ACTION, DEAL, POLICE, FORMER, DANGER, DESIRE, SENATE, RISING, WEIGHT, KNOW, GROUP, RESULT, WINE, BACK, STATUS, COOL

Word list 7: NECK, ESCAPE, TERM, NAME, COFFEE, FIRE, TREE, WALL, LONG, RECORD, SCORE, STORE, CHINA, SHOW, CASE, FORCE, WAIT, TABLE, NATION, COUPLE, MORAL, EVEN, WEST, BEAT, WATCH, BILL, RAIN, POET, LEADER, CHOICE, COVER, PLAY, DOING, LORD, OFFER, REASON, WINTER, DOOR, RADIO, TOUCH, CLASS, PLUS, MARK, BASE, STAGE, TASK, ANSWER, LIFE, BANK, BOTTOM, WISH, RULE, OTHER, DAILY, CLAY, SPOKE, NEED, NIGHT, REGARD, MASTER

Word list 8: LOVE, NUMBER, SQUARE, UNITY, PLANE, DATA, KEEP, FELT, FIGHT, PEACE, DUST, FLOOR, TRIAL, CLEAN, MEAN, EMPTY, FUTURE, FIGURE, COURT, EQUAL, LADY, REST, WRITER, BABY, MANNER, CLUB, SOLID, MODEL, FOOD, RANGE, TASTE, OBJECT, MINE, PERMIT, WOMAN, SHARP, REMOVE, EDITOR, UNIT, HOTEL, POETRY, MIKE, LEAGUE, STORY, HOLE, REPORT, TEXT, TEAM, HILL, COAST, REAL, BRIEF, CAMP, FLOW, SYSTEM, WILL, WIFE, RIVER, FUND, POOL

Supplemental Material: Experiment 1 Congruency Group

Method

Participants. Twenty-four participants (20 female, mean age = 18, SD = 0.82) were recruited from the McMaster University student pool in exchange for course credit. All participants spoke English fluently and had normal or corrected-to-normal vision.

Apparatus and stimuli. The apparatus and stimuli were identical to Experiment 1 with the following exceptions. Stimuli consisted of pairs of red and green interleaved words presented at the centre of the screen. Like in Experiment 1, each word subtended 5.95 degrees of visual angle horizontally, and 0.75 degrees vertically. Together, both words subtended 6.52 degrees horizontally and 1.03 degrees vertically. Examples of the stimuli are presented in Figure S1.

Figure S1: Examples of study phase items for the congruency group. The left side depicts an example of a congruent item; the right side depicts an example of an incongruent item.

Procedure. The procedure was identical to Experiment 1 with the following exceptions. On every study trial, participants saw a white fixation cross in the centre of the screen for 2000 ms, followed by a pair of red and green interleaved words for 1000 ms. Additionally, at the time of test, participants saw a pair of red and green interleaved words, rather than a single red word. It should be noted that old items were presented in the same manner at test as they were at study—that is, old items were presented with the same target-distractor pairing.

Design. The design was identical to Experiment 1 with the following exception. Rather than words being assigned to "repeated", "not-repeated prime", or "not-repeated target" roles, they were assigned to "congruent", "incongruent distractor", or "incongruent target" roles.

Results

Study phase. Correct response times (RTs) were first submitted to the same outlier analysis as Experiment 1 (Van Selst & Jolicoeur, 1994), which eliminated 2.34% of observations from subsequent analyses. Mean RTs were computed from the remaining observations. These mean RTs and corresponding error rates were submitted to two-tailed paired sample t-tests. Naming times were faster for congruent (678 ms) than for incongruent (827 ms) items, $t(23) = 12.6$, $p < .001$, $d = 3.65$, and error rates were lower for congruent (.003) than for incongruent (.022) items, $t(23) = 3.29$, $p = .002$, $d = 0.949$.

Test phase. Congruent and incongruent items were compared using *d'* as a measure of sensitivity and beta as a measure of bias. The mean proportions of "old" responses are displayed in Figure S2.

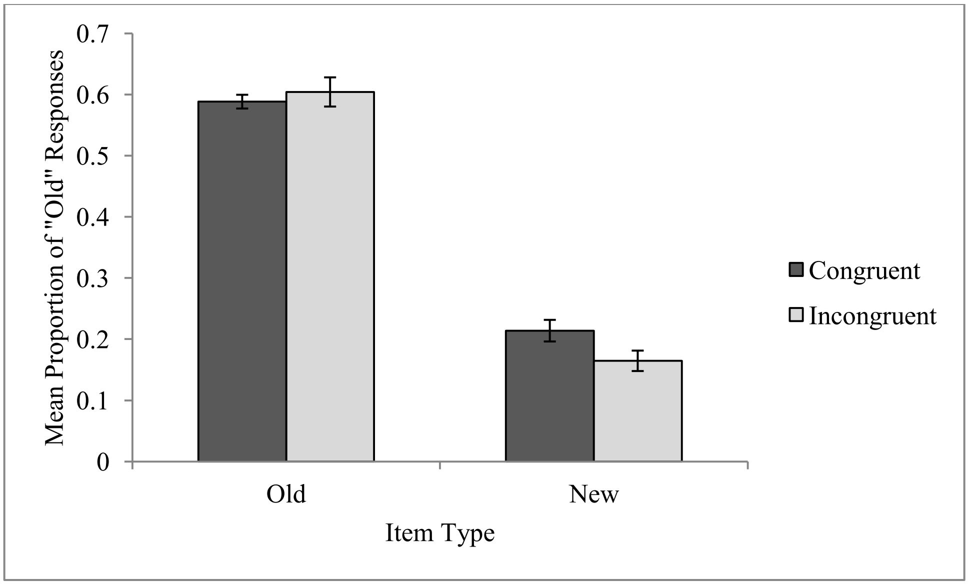

Figure S2: Mean proportions of "old" responses for the congruency experiment. Error bars here represent within-subject error corrected for between subjects variability (Morey, 2008).

Both *d'* and beta were analysed using two-tailed paired sample t-tests. For values of *d'*, the analysis revealed better recognition sensitivity for incongruent (1.37) than congruent items (1.11), $t(23) = 3.27$, $p = .003$, $d = 0.945$. A similar analysis of beta values indicated a criterion difference between incongruent (2.27) and congruent items (1.73) that approached significance, $t(23) = 2.05$, $p = .052$, $d = 0.593$, indicating a more conservative criterion to respond "old" to incongruent than congruent items.

Contributions of recollection and familiarity to memory performance were obtained separately for hit rates and false alarm rates, and the hit minus false alarm difference scores were calculated for each item type. The analysis of recollection and familiarity estimates revealed a pattern similar to Experiment 1. There was no difference in recollection between congruent (.276) and incongruent items (.283), $t < 1$, whereas estimates of familiarity were greater for incongruent (.291) than congruent items (.215), *t*(23) = 2.92, *p* = .008 *d* = 0.843.

Discussion

The present data constitute a nice replication of the congruency experiments reported in Rosner, D'Angelo, MacLellan, and Milliken (2015), in that incongruent items were better remembered than congruent items. The different performance across conditions was restricted to the false alarm rates here, rather than occurring for both hit rates and false alarm rates as in the experiments reported in Rosner, D'Angelo, et al., but this result has a straightforward explanation. In particular, if we assume that congruency at the time of test increases processing fluency, which in turn is attributed to familiarity (Jacoby & Whitehouse, 1989), then the equivalent hit rates for the congruent and incongruent conditions must be a joint product of better recognition of old incongruent than old congruent items, together with a bias to respond "old" to the fluent congruent items. In any case, these results indicate that stimulus congruency has an impact on recognition memory, and that this effect produces results similar to those found when item repetition is manipulated.

It is also worth considering whether the inclusion of distractor words at test in this congruency experiment (but that did not occur in the present repetition experiments) could have contributed to the improved recognition for incongruent items. For example, two different words at study for incongruent items may offer an opportunity for associative encoding between the two words (e.g., the word BRICK interleaved with TRUCK may have led to the image of a truck carrying bricks) that is not available for congruent items. When two words appear at test, an associative cue is then available for incongruent items, but not for congruent items, which might explain the effect that we observed. However, theoretical accounts such as this one were ruled out by Rosner, D'Angelo, et al. (2015) in two ways. First, better recognition for incongruent than congruent study items was also demonstrated in an experiment that presented only single words at test (Experiment 2). Second, recognition was shown to be superior for an incongruent condition in which two different words were interleaved at study than for an incongruent condition in which two different words were spatially separated (Experiment 3).

CHAPTER 3: The function of (dis)fluency: A potential cue for encoding

Rosner, T. M & Milliken, B. (Submitted).

Journal of Experimental Psychology: Learning, Memory, and Cognition

Manuscript ID: XLM-2018-0186

Preface

Chapter 3 presents the results of five experiments in which the relation between the repetition decrement effect and a repetition-driven false recognition effect is examined. The basic method for all experiments involved the presentation of prime-target pairs during both the study and test phases. In Experiments 1-4, the timing of the prime words was varied to better understand the conditions that produce the repetition decrement and false recognition effects. Across these experiments, it was determined that the repetition decrement effect can be produced with both short-duration and long-duration primes, though the size of the effect is larger when long-duration primes are used. Moreover, the false recognition effect was only produced using short-duration primes. More interesting is that taken together, the results of these experiments indicate that when conditions at test allow for the expression of the false recognition effect, the repetition decrement effect is obscured. It is suggested that these results may be due to the timing of responding at test, with fast fluency-based responding driving the false recognition effect precluding the need for slower retrieval which would allow for the expression of the repetition decrement effect. This hypothesis is supported by a re-analysis of the data in Experiment 4 as well as by experimental manipulation of response speed in Experiment 5. The results from these experiments suggest that both the repetition decrement and false recognition effects may be driven by similar encoding processes, though the way retrieval occurs at test determines which of the two effects is observed.

Abstract

Recent studies have demonstrated that immediate repetition of words during a study phase can weaken performance in a subsequent recognition memory test. This *repetition decrement effect* was examined in tandem with the impact of repetition at test on recognition, which can produce a phenomenon known as false recognition—higher false alarm rates for repeated than non-repeated items. We propose that these two effects may be driven by similar processes. Repetition-driven fluency during a study phase may signal that an item is known, which in turn may reduce resources allocated to encoding that item. Repetition-driven fluency at test may be attributed to prior experience with that item, which in turn increases "old" responses to that item regardless of its actual study status. Across five experiments, the false recognition effect appeared to be dependent on rapidly accessed processing fluency, whereas the repetition decrement effect was driven by slower accessed retrieval of the study episode. The two effects were not independent of each other, with the occurrence of the repetition decrement effect being pre-empted by processes that produced the false recognition effect. Nonetheless, the repetition decrement and false recognition effects may both depend on an automatic response to processing fluency/disfluency, with differences between the two effects related to task differences associated with the study and test phases.

Introduction

There has long been interest among memory researchers on the effect of stimulus repetition on memory performance (e.g., Hintzman, 1974; Ratcliff, Sheu, & Gronlund, 1992). Many prior studies have demonstrated that item repetition during a study phase can improve performance on long-term memory tasks[5] (e.g., see Ratcliff, 1990; Stretch & Wixted, 1998). This result has been referred to as a memory strength effect. Studies of the memory strength effect have played an important role in theorizing about recognition memory (e.g., Bruno, Higham, & Perfect, 2009; Jacoby & Dallas, 1981; Stretch & Wixted, 1998; Yonelinas, 1994).

Another repetition effect of interest to memory researchers is the spacing effect (R. A. Bjork & Allen, 1970; Hintzman, 1974). The spacing effect is observed in studies that manipulate spacing of repeated items within a single study session. Memory performance tends to be better for repeated items that are spaced (i.e., presented with other items between the repetitions) than for repeated items that that are massed (i.e., immediate repetitions). This effect is a robust one, and has been observed in recognition (Braun & Rubin, 1998), free recall (Braun & Rubin, 1998; Kahana & Howard, 2005; Underwood, 1969), and cued recall (Cuddy & Jacoby, 1982; Jacoby, 1978). Similar results are found in spacing effect studies that manipulate the study sessions themselves, with better final performance for groups that engage in study sessions across different days than those that mass their study within a single day (Baddeley & Longman, 1978;

---

[5] There are two well-documented effects demonstrating a decrement in performance due to repetition: the Ranschburg Effect (Crowder & Melton, 1965; Jahnke, 1969b; Lee, 1976) and repetition blindness (Kanwisher, 1987). However, these effects concern performance on working memory or immediate recall tasks, rather than delayed memory performance, which is the issue being discussed here.

Bahrick, 1979; R. A. Bjork, 1994). Moreover, the benefits of spaced study sessions have been shown to last over a period as long as eight years (Bahrick & Phelps, 1987). The spacing effect implicates a process that undermines a repeated encoding benefit (i.e., memory strength effect) when spacing between repetitions is short.

The Repetition Decrement Effect

A recent set of studies offers a method with which diminished encoding of repetitions can be examined systematically (Collins et al., 2018; Rosner, López-Benítez, D'Angelo, Thomson, & Milliken, 2018). In these studies, participants were shown prime-target pairs in a study phase, and were asked to read only the targets aloud. Half of the targets matched the prime (repeated) whereas the other half were different from the prime (non-repeated). Participants were then given a surprise recognition memory test in which they responded to single words as "old" or "new". Recognition sensitivity was higher for non-repeated than repeated items. The finding that immediate repetitions were recognized poorly is consistent with deficient processing ideas in the spacing effect literature (e.g., Greene, 1989; Hintzman, 1974). However, the unique property of this result is that immediate repetitions were compared to immediate *alternations* rather than spaced repetitions; that is, better recognition was observed for items seen once than for items seen twice in succession. We call this the repetition decrement effect.

The repetition decrement effect is a potentially useful new tool to study mechanisms that produce diminished encoding of repeated items. One theoretical account of diminished encoding of repetitions was of particular interest to us. This account assumes that the processing consequences of stimulus repetition provide a signal that determines subsequent encoding. For example, repeated items at short spacing may be processed fluently and thereby engender a feeling of familiarity, in effect indicating that the item is already well encoded, and reducing subsequent encoding effort for that item (Greene, 1989; see also E. L. Bjork & Bjork, 2011; R. A. Bjork, 1994). Our interest in this idea was driven by the broad utility of processing fluency cues in cognitive theory, but in particular by processing fluency accounts of false recognition effects (Jacoby & Whitehouse, 1989) that also use a stimulus repetition method.

The False Recognition Effect

In their seminal study, Jacoby and Whitehouse (1989) used a stimulus repetition method to study false recognition effects. Participants were asked to study a list of words for an upcoming memory test. At the time of test, masked prime words were shown just before the target items, and participants judged the target items as “old” or “new”. Some of these target words were the same as the primes (repeated) and some targets were different from the primes (non-repeated). In addition, the primes were presented for a short duration (50 ms) or a long duration (200 ms). In the short prime condition, participants produced higher hits and false alarms for repeated than non-repeated items; this finding was not observed in the long prime condition (possibly due to awareness of the prime word, but see Gellatly, Banton, & Woods, 1995; Joordens & Merikle, 1992). Jacoby and Whitehouse (1989) proposed that repetition increased processing fluency for the target word, and that when participants were unaware of the primes this fluency was attributed to familiarity, thus increasing the likelihood of an “old” response. This systematic tendency to claim to recognize a new item as old is referred to hereafter in this article as the false recognition effect (Jacoby & Whitehouse, 1989; Westerman, 2008).

In principle, the processing fluency cue used to explain the false recognition effect (Jacoby & Whitehouse, 1989) could be adapted to account for diminished encoding observed in studies of the repetition decrement effect (Rosner et al., 2018). Processing fluency attributed to familiarity might well directly influence old/new judgments at the time of test, whereas it could conceivably influence effort automatically allocated to subsequent encoding at the time of study. If this were the case, it would provide a parsimonious account of both effects. However, there are at least two salient differences between these two effects likely to challenge any simple view that they are driven by an identical cue. First, the false recognition effect is a bias effect; that is, stimulus repetition at test increases false alarms, but tends to do the same for hits, leaving the difference between hit and false alarm rates unaffected (e.g., Jacoby & Whitehouse, 1989). In contrast, the repetition decrement effect is a sensitivity effect, with stimulus repetition significantly impacting the difference between hit and false alarm rates. Second, the false recognition effect is typically measured with brief duration masked primes, and indeed often does not occur with longer duration primes (Jacoby & Whitehouse, 1989; but see Bernstein & Welch, 1991; Gellatly et al., 1995; Merikle & Joordens, 1997). In contrast, the repetition decrement effect has only ever been measured with longer duration (e.g., 500 ms) unmasked primes (Collins et al., 2018; Rosner et al., 2018). This discrepancy in the procedure typically used to measure each effect stands in the way of any theoretical account that aims to encompass both effects.

The Present Study

In the present study, the broad goal was to understand the relation between processes that drive the false recognition and repetition decrement effects. Our empirical strategy began with a systematic study of prime duration influences on the two effects when measured in the same task. For all but one of the experiments reported here, we used a prime-target procedure during both a study phase in which participants simply named a set of words, and a recognition test phase in which participants made old/new judgments.

Experiment 1

This experiment provided a first opportunity to measure the influence of immediate repetition both at study and at test on recognition memory. We adapted the method of Rosner et al. (2018) for this purpose by presenting primes immediately prior to targets during both the study and test phases. Primes were of the same duration (500 ms) as in the Rosner et al. study, and therefore we expected to replicate the repetition decrement effect: higher sensitivity for non-repeated than repeated items. Our prediction for the effect of repetition during the test phase was less clear. Although the false recognition effect (i.e., higher false alarms for repeated than non-repeated items) generally occurs only for brief duration masked primes (Jacoby & Whitehouse, 1989; Joordens & Merikle, 1992), a false recognition effect has been reported in studies with longer duration primes that were ignored (Bernstein & Welch, 1991; Merikle & Joordens, 1997) or that were low in salience (Gellatly et al., 1995). Because primes in the present experiment were a different colour than targets (primes were green, targets were red), and because participants were free to ignore the primes, it seemed possible that a false recognition effect would occur. On the other hand, given many prior studies with longer duration primes that have not produced a false recognition effect, it seemed more likely that a false recognition effect would not occur in this experiment.

Method

Participants. Twenty-four participants (18 females, mean age = 20.1) from the McMaster University student pool were recruited in exchange for course credit in introductory psychology. All participants in this and following experiments had normal or corrected-to-normal vision and were fluent English speakers.

Apparatus and stimuli. The experiment was run on a Mac Mini computer with a BenQ LED monitor using PsychoPy software (Peirce, 2007, 2009). The stimuli were presented on a black background in the centre of the screen in Lucida sans console font. The stimulus set consisted of 360 five-letter high-frequency nouns (Appendix A) based on the Kučera and Francis (1967) word corpus. Word frequency ranged from 18-1702 per one million, with a mean of 101. Each word subtended 5.95 degrees of visual angle horizontally, and 0.75 degrees vertically. The letters were 0.66 degrees wide, and were separated by a 0.66 degree space between letters. Prime words were presented in green and target words were presented in red.

Design. The 360 words in the stimulus set were randomly split into six word lists of 60 words each. Word lists were fully counterbalanced across participants, such that each word appeared as “old” or “new” an equal number of times. The lists assigned to be “old” were also counterbalanced across participants, with each list assigned to the roles of “old repeated”, “old non-repeated prime”, and “old non-repeated target” an equal number of times. Note that words assigned to the “repeated” condition served as both prime and target on a trial. The lists assigned to the “new” condition were randomly assigned to “new repeated”, “new non-repeated prime”, and “new non-repeated target” roles. There were 120 trials in the study phase (60 repeated/60 non-repeated) and 240 trials in the test phase (120 old/120 new, 60 repeated/60 non-repeated for both old and new). Finally, “old” items were shown in the exact same manner at study and at test. That is, items that were repeated at study were also repeated at test, and items that were non-repeated at study were non-repeated at test using the same prime-target pairing. Overall, this counterbalancing scheme resulted in twelve possible assignments of word lists to roles, which were used an equal number of times across participants.

Procedure. Participants sat 50 cm in front of the computer monitor. There were three phases in the experiment: a study phase, a distractor phase in which participants completed math problems, and a test phase. Naming times were recorded in the study phase using a microphone attached to a headset worn by participants. Recognition memory decisions were recorded in the test phase as described below.

In the study phase, participants were told that they would see a green word followed by a red word on each trial, and their task was to read aloud the red word as quickly and accurately as possible. They were not informed of an upcoming memory test. Each study phase trial consisted of a white fixation cross for 2000 ms followed by a green prime word for 500 ms. After the green prime, there was a blank screen for 250 ms followed by the red target word for 1000 ms, which participants were to read aloud.

Response times were recorded via microphone for each trial from the time of target onset to the time of voice onset. Following the red target was a blank screen during which the experimenter coded the participants' response accuracy as "correct", "incorrect", or "spoil" using the "1", "2", and "3" keys on the number pad of the keyboard, respectively. A response was considered incorrect if participants said the wrong word, said the wrong word and then corrected their response, or read the word incorrectly. Spoils were recorded when extraneous noise was thought to have triggered the microphone early (e.g., stuttering, coughing). After the study phase, participants completed a 10-minute math distractor task on the computer, which consisted of basic arithmetic problems. Data from the distractor task were not recorded.

Following the distractor phase, participants were given instructions for the recognition memory test phase. They were told they would see a green word followed by a red word on every trial. Their task was to classify the red word as "old" if they remembered it from the study phase and to respond "new" if they did not remember it from the study phase. Participants made "old" responses by pressing the A key on the keyboard and "new" responses by pressing the L key on the keyboard. The trial sequence was similar to the study phase; a white fixation cross appeared for 500 ms, then a green prime for 500 ms, a blank screen for 250 ms, and finally a red target. Together with the target, the words "OLD" and "NEW" appeared on the left and right sides of the screen, respectively, to serve as reminders for which key corresponded to which response. The target and reminder stimuli remained on screen until the participant provided a response.

Following the recognition test, participants were fully debriefed and thanked for their participation.

Results

In this and all following experiments, correct response times (RTs) and naming error rates were the key dependent variables in the study phase. Recognition sensitivity was measured as the difference between hit rate ("old" response to an old item) and false alarm rate ("old" response to a new item). In this and all following experiments, all t-tests were two-tailed paired samples t-tests, and an alpha level of .05 was used for all statistical comparisons.

Study phase. RTs from one participant could not be used due to a microphone malfunction. Correct RTs from the remaining 23 participants were submitted to the non-recursive moving criterion outlier analysis of Van Selst and Jolicoeur (1994), which eliminated 2.5% of RTs from further analysis. The remaining observations were used to compute mean RTs. Means of these mean RTs and corresponding error rates are displayed in Table 1.

The analysis of RTs revealed a significant repetition effect, with faster responses for repeated (507 ms) than non-repeated targets (554 ms), $t(22) = 11.50$, $p < .001$, $d = 2.40$. There was also a significant repetition effect in the analysis of error rates, with a lower proportion of errors for repeated (.007) than non-repeated targets (.016), $t(23) = 2.41$, $p = .024$, $d = 0.49$.

Table 1

Mean response times (ms) and error rates for the study phase

Experiment (Study Phase Prime Duration)	Repeated	Non-Repeated	RT Difference (Non-Rep – Rep)
1 (500 ms)	507 (.007)	554 (.016)	47
2 (50 ms)	515 (.012)	536 (.010)	21
3 (50 ms)	515 (.011)	527 (.015)	12
4 (500 ms)	505 (.010)	544 (.013)	39
5 (500 ms)	501 (.004)	538 (.012)	37

Note: Table displays response times with error rates in parentheses

Test phase. For this and all following experiments, words that were incorrectly named during the study phase were excluded from the test phase analysis. The proportions of "old" responses for each condition of interest were submitted to a 2x2 repeated-measures ANOVA that treated repetition (repeated/non-repeated) and study-status (old/new) as factors. The mean proportion of old responses to old and new items in the repeated and non-repeated conditions are displayed in Figure 1.

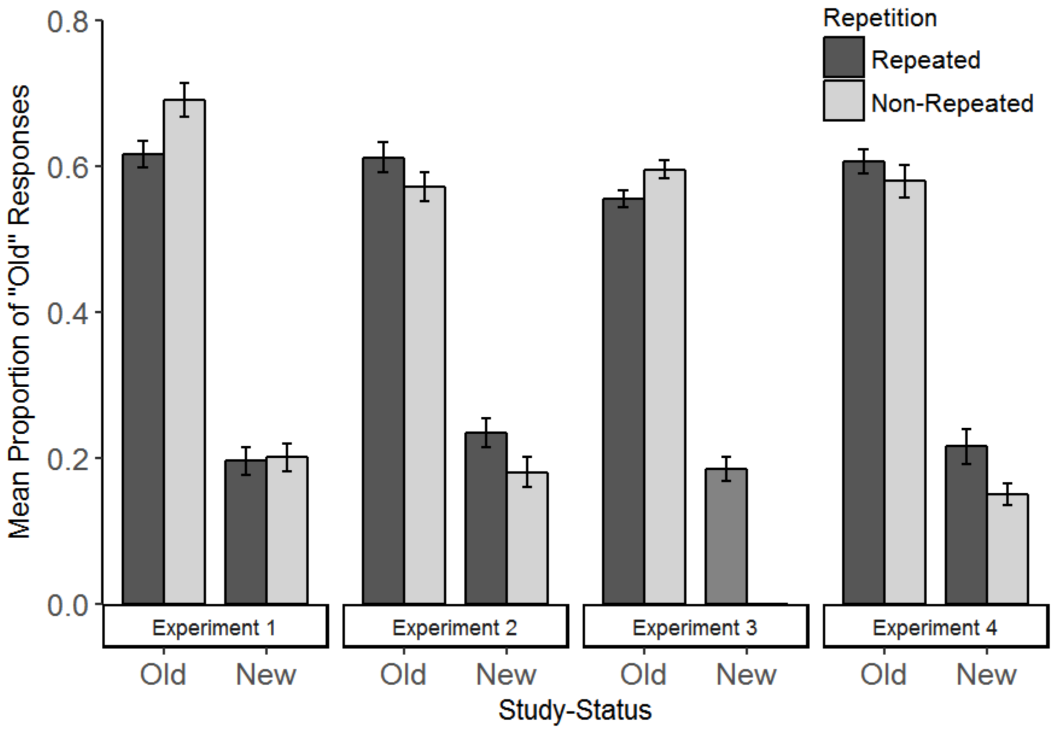

Figure 1: Mean proportions of "old" responses for Experiments 1, 2, 3, and 4. Error bars here and on all other graphs represent within-subject error corrected for between-subjects variability (Morey, 2008).

As the repetition decrement effect is one of higher sensitivity for non-repeated than repeated items, this effect was considered to be observed when there was a greater corrected hit rate (hits – false alarms) for non-repeated than repeated items. On the other hand, the false recognition effect is primarily examined within false alarm rates; as such, this effect was considered to be observed when there was a higher false alarm rate for repeated than non-repeated items.

The analysis revealed a significant main effect of study-status, $F(1,23) = 280.11$, $p < .001$, $\eta_p^2 = .924$, indicating simply that participants were able to discriminate between old (.654) and new items (.198). More important, there was a significant interaction between study-status and repetition, $F(1,23) = 5.77$, $p = .025$, $\eta_p^2 = .201$. This interaction indicates that recognition sensitivity (hits – false alarms) was higher for non-repeated (.491) than repeated items (.419; see Figure 1), and constitutes a replication of the repetition decrement effect reported by Rosner et al. (2018). The interaction was examined further by analyzing the effect of repetition separately for old and new items. For old items, the effect of repetition was significant, $t(23) = 3.07$, $p = .005$, $d = 0.63$, with a higher hit rate for non-repeated (.691) than repeated items (.617). In contrast, for new items, the effect of repetition was not significant, $t(23) < 1$. The absence of a repetition effect in the false alarm rates constitutes a failure to observe the false recognition effect.

Discussion

The results of this experiment were clear. Immediate repetition at study produced higher sensitivity for non-repeated than repeated items. This result constitutes a replication of the repetition decrement effect reported in prior studies (Rosner et al., 2018; see also Collins et al., 2018), and demonstrates that this effect can also be observed when primes are presented at both study and test. However, immediate repetition at test failed to produce a false recognition effect; false alarms were no different for repeated and non-repeated items. This result is in accord with other studies demonstrating that the false recognition effect often depends on brief duration primes, perhaps because brief primes increase the likelihood that perceptual fluency is misattributed to remembering (Jacoby & Whitehouse, 1989).

Experiment 2

In Experiment 2, brief duration (50 ms) masked primes were used rather than longer duration (500 ms) primes both at study and at test. This experiment addressed two important issues. First, it allowed us to examine whether the repetition decrement effect can be measured with brief duration masked primes, an issue not previously addressed in this study or elsewhere. Second, it allowed us to determine whether or not a false recognition effect can be measured with brief duration masked primes in our method (Jacoby & Whitehouse, 1989).

Methods

Participants. Twenty-four participants (17 females, mean age = 19.1) were recruited from the McMaster University student pool in exchange for course credit.

Apparatus, stimuli, design, and procedure. The apparatus, stimuli, design, and procedure were identical to Experiment 1 with the following exceptions. Brief duration primes rather than long duration primes were presented both in the study phase and in the test phase. Primes were pre- and post-masked with a row of five green X's, which superimposed the letter positions of the prime. The primes were presented for 50 ms, and the pre- and post-masks were presented for 500 ms.

Results

Study phase. Correct RTs were submitted to the same outlier analysis as in prior experiments, resulting in the exclusion of 2.1% of the RTs from further analysis. Mean RTs were computed from the remaining observations (Table 1). As in Experiment 1, the effect of repetition was significant, $t(23) = 9.03$, $p < .001$, $d = 1.84$, with faster responses for repeated (515 ms) than non-repeated targets (536 ms). The corresponding effect in the analysis of error rates was not significant, $t(23) < 1$.

Test phase. As in Experiment 1, the mean proportions of "old" responses were submitted to a 2x2 repeated-measures ANOVA, with repetition (repeated/non-repeated) and study-status (old/new) as factors (Figure 1). The analysis revealed a significant main effect of study-status, $F(1,23) = 144.54$, $p < .001$, $\eta_p^2 = .863$, with more "old" responses to old (.592) than new items (.207). More important, there was a significant main effect of repetition, $F(1,23) = 34.24$, $p < .001$, $\eta_p^2 = .598$, with more "old" responses to repeated (.423) than non-repeated items (.375). Also of importance, the interaction between repetition and study-status was clearly not significant, $F(1,23) < 1$. This result indicates that sensitivity (hits – false alarms) was no different for repeated (.379) and non-repeated items (.391) in this experiment.

Although the interaction between repetition and study-status was not significant, a separate contrast of false alarm rates for repeated and non-repeated items was conducted to evaluate whether a false recognition effect occurred. The false alarm rate was indeed higher for repeated (.234) than non-repeated items (.180), $t(23) = 3.87$, $p < .001$, $d = 0.79$.

Discussion

The results of Experiment 2 confirmed that a false recognition effect can be measured with brief duration masked primes in our method. However, there was no evidence of a repetition decrement effect. The absence of a repetition decrement effect implies either that longer duration primes are necessary for the repetition decrement effect or that the influence of short-duration primes at test (and thus the presence of a false recognition effect) makes it difficult to measure a repetition decrement effect. These issues were explored in more detail in Experiment 3.

Experiment 3

To address definitively whether the repetition decrement effect can be observed with brief duration masked primes, the present experiment used brief duration masked primes at study but no primes at test. This method eliminates any possible obscuring of the repetition decrement effect by processes responsible for the false recognition effect. In addition, the sample size was doubled from 24 to 48 to ensure that there was sufficient power to detect a smaller effect than is typically observed with longer duration primes (see Table 1).

Method

Participants. Forty-eight participants (30 female, mean age = 20.2) were recruited from the McMaster University student pool in exchange for course credit or $12.

Apparatus, stimuli, design, and procedure. The apparatus, stimuli, design, and procedure were identical to Experiment 2 with the following exceptions. Brief duration masked primes were presented during the study phase only. During the test phase, no prime words were presented—participants made old/new judgments to a lone presented red target word on each trial.

Results

Study phase. RTs for two participants were not analysed due to a microphone malfunction. Correct RTs were submitted to the same outlier analysis used in prior experiments, which eliminated 2.2% of observations from further analysis. The remaining RTs were used to calculate mean RTs, which were submitted to t-tests along with the corresponding error rates (see Table 1). Again, there was a significant repetition effect in the RT analysis, $t(45) = 3.56$, $p < .001$, $d = 0.525$, with faster RTs for repeated (515 ms) than non-repeated (527 ms) items. No difference in error rates was observed, $t(47) = 1.31$, $p = .195$, $d = 0.190$.

Test phase. As primes were not presented during the test phase, new items could not meaningfully be classified as "repeated" or "non-repeated", and therefore a single false alarm rate was computed (see Figure 1). In a first analysis, the overall hit rate (collapsed across repetition) and false alarm rate were compared. Unsurprisingly, there were more "old" responses to old (.575) than new items (.185), $t(47) = 19.35$, $p < .001$, $d = 2.79$. In a second analysis, the hit rates for repeated and non-repeated items were compared. Importantly, the hit rate was higher for non-repeated items (.595) than for repeated items (.555), $t(47) = 3.10$, $p = .003$, $d = 0.447$.

Discussion

The results of Experiment 3 demonstrate for the first time that a repetition decrement effect can be observed with brief duration masked primes. This finding is particularly important in the present context as it implies that both the false recognition effect and the repetition decrement effect can occur with brief duration masked primes, and therefore that similar processes could, in principle, drive these two effects. Put simply, prime-target repetition may lead to a feeling of familiarity, which at test may elevate false alarms, and at study may reduce encoding effort and thus lower sensitivity on a following recognition test.

At the same time, an issue that remains to be explained is why a repetition decrement effect was observed in Experiment 3 with brief duration masked primes, and yet there was no evidence for such an effect in Experiment 2 with the same type of primes. We addressed this issue in the final two experiments.

Experiment 4

To this point, a repetition decrement effect was observed in Experiments 1 and 3, and a false recognition effect was observed in neither experiment. Further, a repetition decrement effect was not observed in Experiment 2, and this was precisely the experiment in which a false recognition effect was observed. Together, these results suggest that processes that drive the false recognition effect at test may obscure the repetition decrement effect in some manner. As a strong test of this idea, the present experiment used longer duration (500 ms) primes at study and brief duration masked (50 ms) primes at test. The rationale for this experiment was that longer duration primes at study appear to produce a more robust repetition decrement effect than brief duration masked primes (see Experiments 1 and 3). If a repetition decrement effect is not observed with long duration primes at study and brief duration masked primes at test, then this result would converge with Experiment 2 in demonstrating that brief primes at test influence performance in a way that obscures the repetition decrement effect.

Method

Participants. Thirty-six participants (27 females, mean age = 18.4) were recruited from the McMaster University student pool in exchange for course credit.

Apparatus, stimuli, design, and procedure. The apparatus, stimuli, design, and procedure were identical to prior experiments with the following exception. During the study phase longer duration primes (500 ms) preceded the targets, whereas during the test phase brief duration masked primes (50 ms) preceded the targets.

Results

Study phase. The same outlier analysis used in prior experiments eliminated 2.3% of RTs from further analysis. Mean RTs computed from the remaining observations and corresponding error rates were then subjected to paired sample t-tests (see Table 1). A significant repetition effect was observed in the RTs, $t(35) = 11.08$, $p < .001$, $d = 1.85$, with faster RTs for repeated (505 ms) than non-repeated (544 ms) items. The difference in error rates was not significant, $t(35) = 1.00$, $p = .324$, $d = 0.167$.

Test phase. Proportion "old" responses in each condition were submitted to a 2x2 repeated-measures ANOVA that treated repetition (repeated/non-repeated) and study-status (old/new) as factors (see Figure 1). A significant main effect of study-status was observed, $F(1,35) = 311.04$, $p < .001$, $\eta_p^2 = .899$, simply indicating that participants were able to discriminate between old (.593) and new items (.183). There was also a significant main effect of repetition, $F(1,35) = 4.52$, $p = .041$, $\eta_p^2 = .114$, with more "old" responses to repeated (.411) than non-repeated items (.365). Critically, the interaction between repetition and study-status was not significant, $F(1,35) = 2.53$, $p = .120$, $\eta_p^2 =$

.067. The non-significant interaction indicates that recognition sensitivity (hits – false alarms) was not significantly higher for non-repeated (.429) than repeated items (.391), which constitutes a failure to replicate the repetition decrement effect reported in Experiment 1. However, a separate analysis of false alarm rates revealed that a false recognition effect was observed. The false alarm rate was higher for repeated (.215) than for non-repeated items (.150), $t(35) = 2.58$, $p = .014$, $d = 0.43$.

Discussion

The results of this experiment revealed a significant false recognition effect and a non-significant repetition decrement effect. This pattern of results is consistent with that observed in Experiment 2, and again suggests that the false recognition effect obscures in some way the repetition decrement effect. The repetition decrement effect was robust in Experiments 1 and 3 when the false recognition effect was not present, but absent in Experiments 2 and 4 when the false recognition effect was present.

Two Bases of Responding

The results of Experiments 2 and 4 failed to demonstrate a repetition decrement effect under conditions in which a false recognition effect was present. These results imply a dependent relation between these two effects–the presence of the repetition decrement effect may depend on whether or not performance at test is driven by processes that produce the false recognition effect. The false recognition effect is commonly thought to occur because stimulus repetition at test eases processing (Jacoby & Dallas, 1981), and this processing ease (or fluency) is misattributed to prior experience (Jacoby & Whitehouse, 1989; Westerman, 2008; Whittlesea, 1993). This fluency misattribution is also thought to hinge on automatic processes available quickly after target onset. Indeed, prior studies have demonstrated that forcing participants to respond quickly in a recognition task leads to an increase in fluency-based responding (Boldini, Russo, & Avons, 2004; Boldini, Russo, Punia, & Avons, 2007; Espinosa-García, Vaquero, Milliken, & Tudela, 2017; Parks, 2013) . In contrast, retrieval processes that involve a deliberate search of memory unfold more slowly (Jacoby & Dallas, 1981; Quamme, Yonelinas, & Norman, 2007; Rotello, Macmillan, & Van Tassel, 2000; but see Mulligan & Hirshman, 1995).

With this framework in mind, it seemed possible that fast, fluency-based responding (due to the effect of repetition at test) mediates the influence on performance of slower-developing retrieval processes that drive the repetition decrement effect. If this conjecture is correct, then these two effects may vary systematically as a function of the speed of recognition responses. This basic premise has proved useful in prior studies of recognition. For example, modality match effects (which are based on overlap in perceptual processing) have been shown to be larger for fast than slow recognition responses (Boldini et al., 2004; Parks, 2013). In contrast, in the same studies, levels of processing effects (which are based on slower to access conceptual information) were shown to be larger for slow than fast recognition responses (Boldini et al., 2004; Parks, 2013). Similarly, it has been found that speed of responding can impact whether or not the picture-superiority effect is observed (Boldini et al., 2007), with evidence for fluency-based responding when recognition decisions are made quickly, and more retrieval-based responding when recognition decisions are made more slowly.

By extension, fast recognition responses may drive the false recognition effect, whereas slower recognition responses may underlie the repetition decrement effect. To examine this hypothesis, an additional analysis was conducted on the results of Experiment 4. Specifically, the recognition data were separated by RT quartile; that is, the fastest 25% of responses at test was separated from the next fastest 25% of responses, and so on[6]. The RTs were first submitted to a one-way repeated-measures ANOVA that treated quartile (first/second/third/fourth) as the only factor. Unsurprisingly, there was a significant difference in RTs across quartiles, $F(3, 105) = 172.32$, $p < .001$, $\eta_p^2 = .871$. The mean RT was 644 ms (SD = 135 ms) in quartile 1; 874 ms (SD = 214 ms) in quartile 2; 1210 ms (SD = 367 ms) in quartile 3; and 2671 ms (SD = 1001 ms) in quartile 4.

The mean proportion of "old" responses in each quartile are displayed in Figure 2. These proportion "old" judgments were submitted to a 4x2x2 repeated-measures ANOVA, with quartile (first/second/third/fourth), repetition (repeated/non-repeated), and study-status (old/new) as factors. The 3-way interaction approached significance, $F(3,105) = 2.23$, $p = .088$, $\eta_p^2 =.162$, and was examined further by conducting separate 2x2 ANOVAs for each quartile. All of these ANOVAs revealed a main effect of study-status (all *p*'s < .001), indicating better than chance recognition sensitivity. Other findings are noted separately for each quartile below.

---

[6] The analyses presented did not exclude outlier RTs at test; that is, decisions that were made both unusually slowly and unusually quickly were included. The reason for this decision was twofold. First, clipping off the extreme responses would reduce the range of responses analysed. Removing outliers may especially impact the first and fourth quartiles, which was not our intention. Second, separating responses by quartile already reduced our cell size, with a maximum of 60 responses in each quartile per participant; removing the outliers would lead to a further reduction of cell size, and thus reduce statistical power. Finally, it should be noted that the same analyses were conducted with outliers excluded. Doing so led to the same general pattern of data for each quartile, although the 3-way interaction no longer approached significance (likely due to reduced power, as mentioned above).

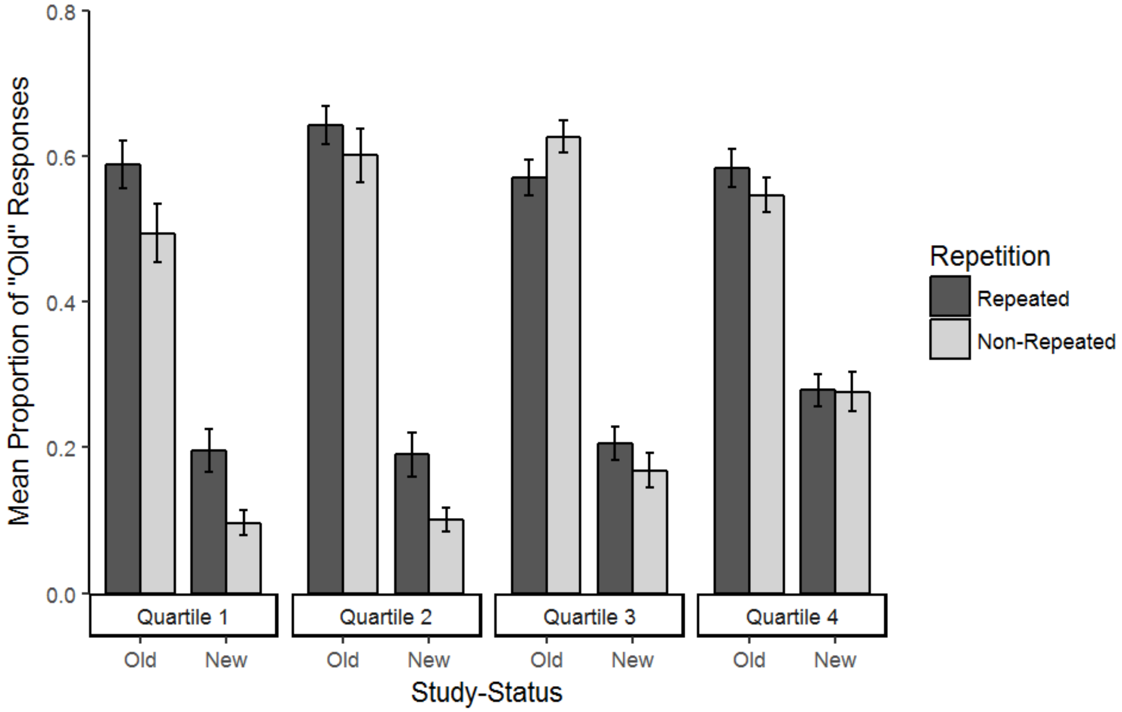

Figure 2: Mean proportions of “old” responses for the four RT quartiles in Experiment 4.

In quartile 1, the main effect of repetition was significant, *F*(1,35) = 10.49, *p* = .003, η_p^2 = .231, with more “old” responses to repeated (.391) than non-repeated items (.295). The interaction between repetition and study-status did not approach significance, *F* < 1. A separate analysis of false alarms revealed a significantly higher false alarm rate for repeated (.195) than for non-repeated items (.097), *t*(35) = 2.95, *p* = .006, *d* = .492. In summary, in quartile 1 there was a false recognition effect but no evidence of a repetition decrement effect.

In quartile 2, the main effect of repetition approached significance, *F*(1,35) = 3.69, *p* = .063, η_p^2 = .095, with more “old” responses to repeated (.416) than non-repeated (.351) items. Again, the interaction between repetition and study-status was not significant, $F(1,35) = 1.62$, $p = .211$, $\eta_p^2 = .044$. A separate analysis of the false alarms revealed a higher false alarm rate for repeated items (.190) than non-repeated items (.100), $t(35) = 2.64$, $p = .012$, $d = .441$. In summary, as for quartile 1, in quartile 2 there was a significant false recognition effect but no repetition decrement effect.

In quartile 3, the main effect of repetition was not significant, $F < 1$. However, the interaction between repetition and study-status was significant, $F(1,35) = 5.79$, $p = .021$, $\eta_p^2 = .142$, with higher sensitivity for non-repeated than repeated items (corrected hits were .458 and .365, respectively). The simple main effects of repetition were non-significant for both old and new items (all *p*'s > .10). In summary, in contrast to quartiles 1 and 2, in quartile 3 there was a significant repetition decrement effect but no significant false recognition effect.

Finally, in quartile 4, neither the main effect of repetition nor the interaction between repetition and study-status were significant, *F*'s < 1. A separate t-test of the repetition effect for false alarms also failed to reveal a significant effect, $t < 1$. These analyses indicate that in quartile 4 there were no reliable effects of repetition on performance.

The quartile analysis indicates that the false recognition effect was present with fast responding and diminished as recognition decisions were made more slowly. In addition, the repetition decrement effect emerged only for the slower quartile 3 responses, at which point the false recognition effect was no longer present. These results are largely consistent with the hypotheses outlined above: that the false recognition effect is driven by relatively fast access to fluency whereas the repetition decrement effect occurs when fast recognition decisions fail to pre-empt the slower retrieval processes upon which the repetition decrement effect is based. The interested reader will find corresponding quartile analyses for Experiments 1, 2, and 3 in Appendix B, which produced results that are broadly consistent with those of Experiment 4.

Experiment 5

The goal of Experiment 5 was to examine the dependence of the false recognition and repetition decrement effects on response speed using a response deadline procedure. The results of prior studies using response deadline procedures suggest that forcing fast responding leads to a reliance on fluency, whereas forcing slower responding leads to a reliance on slower retrieval processes (Boldini et al, 2004, 2007; Espinosa-García et al., 2017; Parks, 2013; but see also Mulligan & Hirshman, 1995). If the false recognition effect is driven by fast, fluency-based processes and the repetition decrement effect is driven by controlled retrieval, then a response deadline procedure should tease these two effects apart. Specifically, participants forced to provide recognition decisions within a few hundred milliseconds should produce only the false recognition effect, whereas participants forced to wait more than a second to provide a recognition decision may produce only the repetition decrement effect.

Method

Participants. Forty-eight participants (45 female, mean age = 18.5) were recruited from the McMaster University student pool. Twenty-four participants were randomly assigned to each of the short-deadline and long-deadline conditions.

Apparatus, stimuli, design, and procedure. The apparatus, stimuli, design, and procedure were identical to Experiment 4 with the following exceptions. The test phase instructions included a description of the response deadline procedure (see below). After the test phase instructions but prior to the test phase, participants completed a training phase to get accustomed to the response deadline procedure. Each training phase trial consisted of a fixation cross for 500 ms, followed by a green pre-mask for 500 ms, a random string of five letters in green for 50 ms, a green post mask for 500 ms, and then finally a random string of five letters in red; items were repeated or non-repeated to mimic the design of the actual test phase. Following the onset of the red letter string, a row of white asterisks appeared underneath the word after either 250 ms or 1500 ms for the short- and long-deadline conditions, respectively. Participants were told that the row of asterisks was their cue to provide a response, and they were to do so within 400 ms of the onset of the asterisks. Responses provided after 400 ms produced a beep to let them know they should respond faster. They were also told not to respond prior to the onset of the cue; if they did, the message, “Please wait for the asterisks to provide your response,” was shown on screen for two seconds after the trial. In total, participants in the short-deadline condition had 650 ms from target onset to respond, whereas those in the long-deadline condition had 1900 ms from target onset to respond. For the training trials, participants were told to decide whether or not the letter “T” was present in the red letter string; they were to press the A key for “yes” and the L key for “no”. The stimuli disappeared after participants responded. Participants were required to complete a minimum of 20 training trials. The training phase ended when participants were confident in their ability to complete the task. On average, participants in the short-deadline condition completed 28 training trials and participants in the long-deadline condition completed 27 training trials.

Following completion of the training trials, participants were reminded of the instructions for the recognition memory test. They were told that it would be the same as the training phase they just completed, but that they were to identify whether or not the red word was old. They were to respond "yes" by pressing the A key if they thought the word was old, and "no" by pressing the L key if they thought the word was new. The timing of the test phase was identical to that of the training phase.

Results

Study phase. RTs were submitted to the same outlier analysis as prior experiments, resulting in the exclusion of 2.3% of observations from further analysis. Mean RTs and error rates were submitted to 2x2 mixed-factor ANOVAs, with repetition (repeated/non-repeated) as a within-subject factor and deadline condition (short-deadline/long-deadline) as a between-subjects factor (see Table 1). The main effect of repetition was significant in both the analysis of RTs, $F(1,46) = 178.62$, $p < .001$, $\eta_p^2 = .795$, and error rates, $F(1,46) = 5.59$, $p = .022$, $\eta_p^2 = .108$. Consistent with prior experiments, repeated items were responded to faster (501 ms) and produced fewer errors (.004) than non-repeated items (538 ms; .012). As the study phases were identical for the two groups, it was unsurprising that no other effects were significant, all *p*'s > .10.

Test phase. Proportion "old" responses were submitted to a 2x2x2 mixed-factor ANOVA, with repetition (repeated/non-repeated) and study-status (old/new) as within-subject factors and deadline condition (short-deadline/long-deadline) as a between-subjects factor. These data are presented in Figure 3. The main effect of study-status was significant, $F(1,46) = 247.21$, $p < .001$, $\eta_p^2 = .843$, with more old responses to old (.480) than new items (.243). In addition, all 2-way interactions were significant, as described below.

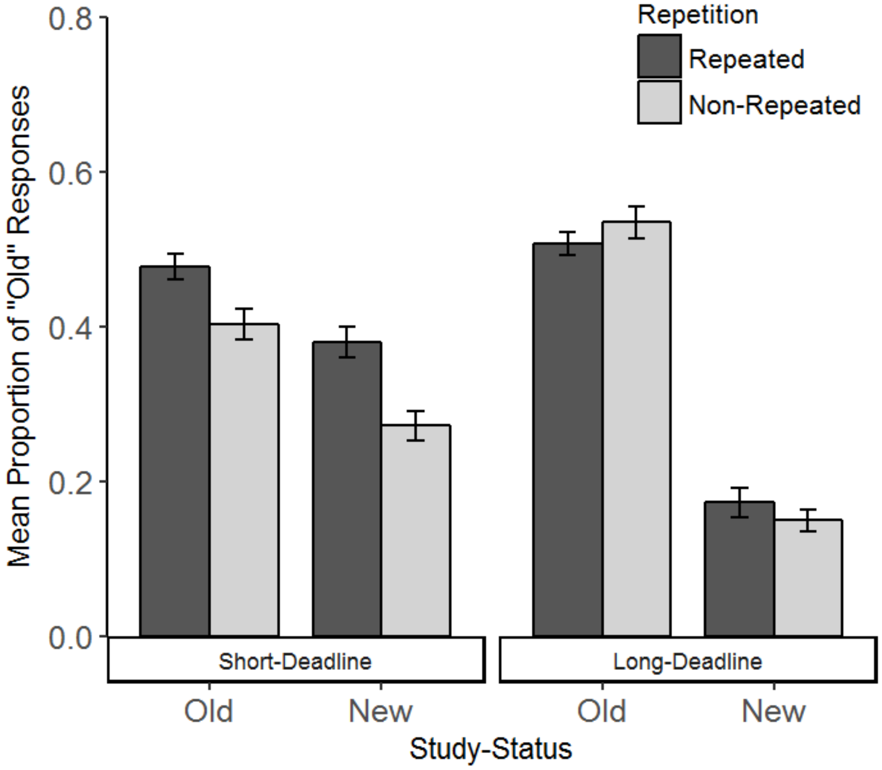

Figure 3: Mean proportions of “old” responses for Experiment 5.

The deadline condition by study-status interaction was significant, $F(1,46) = 66.51$, $p < .001$, $\eta_p^2 = .591$. Although both groups had higher hit than false alarm rates, this difference (and therefore recognition sensitivity) was greater for the long-deadline condition (.360) than the short-deadline condition (.114). The deadline condition by repetition interaction was also significant, *F*(1,46) = 12.05, *p* = .001, η_p^2 = .208, and indicated that the effect of repetition (regardless of old/new status) differed between groups. The simple main effect of repetition for the short-deadline condition was significant, *t*(23) = 4.04, *p* < .001, *d* = 0.825, indicating more "old" responses to repeated (.428) than non-repeated items (.337). The simple main effect of repetition was not significant for the long-deadline condition, *t* < 1. This result is consistent with the idea that the false recognition effect should be observed for fast responses but perhaps not for slower responses. Finally, a significant repetition by study-status interaction, *F*(1,46) = 5.34, *p* = .025, η_p^2 = .104, indicated that corrected hits (hits – false alarms) differed for repeated and non-repeated items. Close inspection of Figure 1 indicates that sensitivity was higher for non-repeated than repeated items (corrected hits were .258 and .216, respectively). This result constitutes evidence of a repetition decrement effect in the data collapsed across the two deadline conditions.

Although the three-way interaction between deadline condition, repetition and study-status was not significant, *F* < 1, the most important analyses in this experiment relate to our *a priori* predictions for the two deadline conditions. As such, separate 2x2 ANOVAs were conducted for each deadline condition with repetition and study-status as within-subject factors.

For the short-deadline condition, there was a significant main effect of study-status, *F*(1,23) = 37.2, *p* < .001, η_p^2 = .618, with more "old" responses to old (.440) than new items (.326). There was also a significant main effect of repetition, *F*(1,23) = 16.3, *p*

$< .001$, $\eta_p^2 = .415$, with more “old” responses to repeated (.428) than non-repeated items (.337). A separate analysis of the false alarms also revealed a higher false alarm rate for repeated than non-repeated items, $t(23) = 3.91$, $p < .001$, $d = .799$. This result illustrates a false recognition effect in the short-deadline condition. Importantly, the interaction between repetition and study-status was not significant, $F(1,23) = 1.40$, $p = .248$, $\eta_p^2 = .057$, indicating that a repetition decrement effect was not observed in the short-deadline condition.

In the long-deadline condition, there was a main effect of study-status, $F(1,23) = 231.53$, $p < .001$, $\eta_p^2 = .910$, again indicating more “old” responses for old (.521) than new items (.161). Most important, the interaction between repetition and study-status was significant, $F(1,23) = 5.07$, $p = .034$, $\eta_p^2 = .181$, indicating greater sensitivity for non-repeated than repeated items (corrected hit rates were .385 and .335, respectively). A separate analysis of the false alarms revealed no significant difference between the repeated and non-repeated conditions, $p > .10$. To summarize, in the long-deadline condition there was higher sensitivity for non-repeated than repeated items (i.e., a repetition decrement effect) and no false recognition effect.

Discussion

The results of this experiment were perfectly in line with our predictions and with the RT quartile analysis of Experiment 4. In the short-deadline condition, a repetition decrement effect was not observed, and a false recognition effect was observed. In the long-deadline condition, a repetition decrement effect was observed, and a false recognition effect was not observed. Notably, the participants in both groups had identical encoding phases, so the difference in performance must be due to how the response deadline manipulation influenced retrieval. These results support the proposal that the false recognition effect is driven by quickly accessed information, whereas the repetition decrement effect is driven by more slowly accessed information. The implications of this proposal for our broader understanding of the repetition decrement and false recognition effects are addressed in more detail in the General Discussion.

General Discussion

The goal of the current study was to understand how two effects of repetition on recognition memory—the repetition decrement and the false recognition effect—are related. Both effects are the product of immediate repetition; at study, immediate repetition leads to lower recognition sensitivity than immediate alternation (the repetition decrement effect), and at test, immediate repetition leads to more "old" responses than immediate alternation (the false recognition effect). The five presented experiments paint a clear picture: although these two effects are elicited by a similar method, they appear to be driven by different retrieval processes. Namely, the false recognition effect owes to fast-acting fluency-driven responses at test, whereas the repetition decrement effect emerges when responding is slower, and thus items that were better encoded at study have a chance to be retrieved at test. Moreover, the presence of the repetition decrement effect appears dependent on whether or not a false recognition effect is produced. In each of Experiments 1 and 3, a repetition decrement effect was observed. In each of Experiments 2 and 4 the same manipulations were used at encoding, and yet the repetition decrement effect was eliminated and a false recognition effect was observed.

Moreover, the results of both the quartile analysis of Experiment 4 (see also the quartile analyses for Experiments 1-3 in Appendix B) and Experiment 5 support the notion that if responses are driven by fast-acting fluency, the repetition decrement effect will be difficult to measure.

Disfluency and the False Recognition Effect

The original goal identified for this set of studies was to examine the relation between the repetition decrement and false recognition effects. Specifically, we were interested in the idea that both effects could be related to processing fluency for repeated items; processing fluency at study might signal that additional encoding is not needed, whereas processing fluency at test might increase the likelihood of an "old" judgment (a misattribution of fluency to familiarity). Yet, the results of our empirical work point to these two effects being driven by different retrieval processes, with the repetition decrement effect related to retrieval that unfolds relatively slowly, and the false recognition effect related to quick access to processing fluency. In light of these results, it is worth re-considering whether a framework centered on processing fluency can accommodate both of these results. Indeed, we propose that such a framework exists, but that it hinges centrally not on fluency, but *dis*fluency as a signal that drives both encoding at study and recognition decisions at test. Moreover, this framework is not incompatible with the notion of the two effects being driven by different retrieval processes at test.

We propose that disfluent processing of non-repeated items may result in better encoding of those items. This suggestion is in line with previous work demonstrating that disfluency may trigger automatic higher-order encoding processes that allow us to learn information that is otherwise more difficult to understand (Alter, Oppenheimer, Epley, & Eyre, 2007). When experienced at study, this disfluency and resulting increase in encoding leads to the repetition decrement effect measured at test (Experiments 1, 3, and 5).

In the context of the current study, this automatic increase of higher-order encoding in response to disfluency would also be expected to occur at test, with non-repetitions being better encoded than repetitions. This disfluency and resulting encoding can also account for the observed false recognition effect (Experiments 2, 4, and 5). Classic accounts of false recognition propose that fluency is misattributed to familiarity, which increases false alarm rates for repeated relative to non-repeated items (e.g., Jacoby & Whitehouse, 1989; Whittlesea, 1993). However, the focus on processing fluency for repeated items, rather than processing disfluency for non-repeated items, is an arbitrary distinction–presumably it is the processing fluency/disfluency of one item type relative to the other that serves as the basis for attributions of familiarity/unfamiliarity to test items (Jacoby & Dallas, 1981; Westerman, 2008; Whittlesea & Williams, 1998). In other words, processing disfluency for non-repeated items at test could provide a signal that is used to judge that an item is new rather than old. Returning to the current set of experiments, we propose that the signal to which participants may have access is the automatic engagement of higher level encoding that occurs for non-repeated relative to repeated items. In effect, at the time of test, participants may be sensitive to their own learning; to the extent that higher order learning is automatically triggered at test in response to disfluency, participants may phenomenologically experience that learning as having encountered a new item.

With this framework in mind, the same automatic process in response to disfluency can lead to both the repetition decrement and false recognition effects. At both study and test, disfluent non-repetitions may be better encoded than repetitions. When this process occurs at study, it produces the repetition decrement effect: higher sensitivity for non-repeated than repeated items. When this same encoding process occurs at test, it can serve as a cue that an item is not known, and is observed as the false recognition effect: fewer "old" responses for non-repeated than repeated items. Which of the two effects is observed is dependent on the time-course of responding. To the extent that controlled retrieval is able to occur at test, the repetition decrement effect will be observed. Alternatively, faster recognition responses may be driven by more immediately available cues (i.e., whether or not the test item in question has triggered higher level encoding), thus producing the false recognition effect. If this proposal is correct, and automatic encoding of disfluent test items provides a cue for responding, then a second test phase ought to reveal better recognition memory for non-repeated than repeated items experienced during the initial test phase. This is an issue that is well worth examining in future research.

Perceptual Versus Conceptual Processing

Of course, the above proposal is only one possible explanation of the results presented here. Although we have suggested that both effects are driven by the same process, an alternative account might attribute the two effects to different processes. In particular, prior studies that have used a response deadline procedure suggest that the processes driving fast and slow recognition decisions can be qualitatively different. Whereas modality match effects are largest when responding is fast and diminish with increased time-to-respond, levels of processing effects only appear at longer response deadlines (Boldini et al., 2004; see also Parks, 2013). The results of these studies suggest that the modality match effect depends on fast-acting perceptual processing that may be fluency-driven, whereas other effects (such as levels of processing) are more conceptually-based and involve slower retrieval processes (Boldini et al., 2004; Parks, 2013). On a related theme, application of the response deadline procedure to the picture superiority effect (better memory for items shown as pictures than as words at encoding, despite all items shown as words at test) has revealed that this effect emerges only at longer response deadlines. When response deadlines are short, the picture-superiority effect is reversed; that is, memory is better for items that are perceptually similar at study and test (Boldini et al., 2007). Again, these results demonstrate that perceptually-based effects may be driven by fast-acting fluency, whereas conceptually-based effects emerge only when responding is slower.

Turning to the present results, the same contrast between perceptual and conceptual processing may apply; the repetition decrement effect may be driven by conceptual processing, whereas the false recognition effect may be more perceptually-based. According to this view, conceptual processing may be responsible for better encoding of non-repeated items, possibly triggered by the more difficult processing itself (e.g., R. A. Bjork, 1994). Alternatively, the availability of fast-acting perceptual processing may be the driving force behind the false recognition effect. This proposal is certainly consistent with the results of the current study, and Experiment 5 in particular. However, the results of prior studies using the response deadline method (Boldini et al., 2004, 2007; Parks, 2013) were aimed purposefully at differences in the *kind* of information that was encoded and retrieved, with levels of processing aimed at conceptual processing and modality match/mismatch aimed at perceptual processing. In contrast, our experiments did not aim purposefully at a manipulation of either perceptual or conceptual processing. As such, any inference about perceptual/conceptual processing differences for the false recognition and repetition decrement effects in our experiments hangs on an analogy to prior work using the response deadline procedure. Additional work needs to be done to evaluate whether such an inference can be supported empirically.

The Repetition Decrement Effect and Repetition Suppression

It is worth noting a parallel between the repetition decrement effect and findings of repetition suppression in neuroimaging studies (for reviews, see Gotts, 2016; Henson & Rugg, 2003). Broadly speaking, repetition suppression refers to lowered neural activity that accompanies repeated presentations of stimuli. The parallel to the behavioural results reported here is that the repetition decrement effect might be described as demonstrating a “suppression” of encoding that produces lower hit rates for repetitions than non-repetitions.

Although much research needs to be done to confirm this link, it is plausible that the repetition decrement and repetition suppression effects are related. That is, the lowered neural activity measured in studies of repetition suppression may well be a biological expression of poor item encoding that lowers recognition sensitivity for repeated items. One approach to studying the relation between the repetition decrement and repetition suppression effects would be to capitalize on the established link between the spacing and repetition suppression effects (Xue et al., 2011; Zhao et al., 2015). If all of these effects are fundamentally related, then it ought to be possible to demonstrate a close association between the repetition decrement and spacing effects. Indeed, recent work does point to such an association (Collins & Milliken, 2018). Of course, a more direct approach would aim to examine repetition suppression within the context of the procedure used in the current study. If the repetition decrement effect is accompanied by repetition suppression, it would then be important to examine the causal link: for example, if repetition suppression is somehow disrupted, will the repetition decrement effect also be disrupted?

Conclusions

The findings of the current study suggest that two immediate repetition effects on recognition, one at study and the other at test, can both be measured using the same method, and may reflect similar encoding processes. However, the repetition decrement effect and the false recognition effect ultimately appear to reflect different retrieval processes. Quickly accessed fluency-based processing appears responsible for the false recognition effect, whereas more slowly accessed retrieval processes appear responsible for the repetition decrement effect. Importantly, fluency-based decisions can pre-empt the expression of slower retrieval processes in recognition performance, making the repetition decrement effect difficult to measure when the false recognition effect is also present. Nonetheless, we have proposed that processing (dis)fluency could drive both the decisions that produce the false recognition effect and additional encoding that produces the repetition decrement effect. Further research should aim to test this idea directly, as well as to examine the relation among the repetition decrement, spacing, and repetition suppression effects.

Appendix A

Word List 1: ROUND, VOICE, BIRTH, CHAIR, BOUND, SENSE, DOUBT, MASON, GREEN, STUFF, CRIME, STERN, ENTRY, STAFF, MOVIE, GRACE, SCOPE, MODEL, TOUGH, ANGER, SLIDE, CYCLE, MONEY, PAUSE, BEACH, BREAK, DRINK, CRAFT, STORY, POINT, LUNCH, BASIS, PLAIN, UNDER, COVER, CHIEF, WIDOW, OCEAN, RIVER, LEMON, SPLIT, LOBBY, TRUTH, GROUP, PRIZE, CLOTH, TOUCH, TRUCK, MAJOR, DELAY, MOTOR, SHAPE, CHEEK, GRAIN, WOMAN, DROVE, DRAMA, TITLE, EARTH, SERUM

Word List 2: ORDER, OFFER, SWEET, ENEMY, GROSS, PLANT, FRONT, BRIEF, NOBLE, STUDY, PRIME, THEME, MOUTH, SLEEP, NIGHT, BLAME, ACTOR, SIXTY, FRAME, PAINT, GRANT, BREAD, TREAT, SHELL, CHEST, STOCK, PANIC, DANCE, ASIDE, REALM, TEETH, PLATE, CRASH, DRIFT, SWIFT, CATCH, DOING, CLOSE, PIANO, DRIVE, PHONE, SMILE, WHILE, WATCH, WOUND, MORAL, SHAME, KNIFE, APRIL, COURT, ISSUE, OWNER, FOCUS, CLOUD, CAUSE, HURRY, POUND, CLASS, COACH, IMAGE

Word List 3: OTHER, CHECK, PHASE, GRASS, FLOOR, STEEL, ESSAY, DRAFT, COUNT, MAGIC, LAUGH, STRIP, WRONG, CREAM, TOOTH, SCORE, FLUID, SHORE, THREE, FORCE, SERVE, SMELL, METAL, DEPTH, MATCH, PUPIL, PRIDE, TRAIN, OPERA, LODGE, MERIT, BURST, STAND, QUEEN, LATIN, NOVEL, DAILY, STICK, TRADE, SUGAR, WAGON, HONEY, PITCH, CHOSE, CRACK, WORST, CURVE, LABEL, CROWN, RIGHT, PLANE, TODAY, SMITH, SPACE, THROW, AGENT, IDEAL, RIDGE, DRILL, MUSIC

Word List 4: COAST, SPELL, JOINT, CABIN, QUICK, SOUND, PANEL, SHARE, STATE, STALL, FENCE, DAIRY, SLOPE, TRACK, SHEEP, GIANT, JUDGE, CLOCK, TRAIL, INNER, SPEED, MOTEL, PEACE, SOLID, GLASS, PENNY, START, SHOCK, PROOF, CHILD, CLEAN, SHEAR, SOUTH, WHEEL, THICK, NOISE, SCALE, CROWD, HEART, SEVEN, VALUE, CHAIN, CHASE, INDEX, WASTE, SQUAD, BRASS, SKILL, EXTRA, MINOR, RADIO, GUARD, SHORT, YIELD, DOZEN, HORSE, THIRD, RANGE, GUILT, ERROR

Word List 5: WORLD, PRESS, EVENT, BRAIN, PILOT, LEAST, ONSET, SNAKE, ANGLE, FIELD, GUESS, TASTE, TREND, GLORY, INPUT, SHARP, DRAIN, SMALL, STILL, LOCAL, PRIOR, SAUCE, STONE, RANCH, FAINT, LEVEL, QUIET, FLASH, SWING, MIDST, EMPTY, LIGHT, UPPER, BLOCK, THING, WHITE, GUIDE, MONTH, TRUST, CHARM, ROUGH, CLERK, YOUTH, RATIO, WATER, GUEST, EIGHT, SHADE, PARTY, CHART, FLOOD, SHIFT, SIGHT, BRIDE, LIMIT, EQUAL, SMART, FIGHT, ADULT, ROUTE

Word List 6: PRICE, TABLE, PORCH, NORTH, BRICK, LOOSE, SCENE, MERCY, FIFTY, MOUNT, SHOOT, SKIRT, MARCH, STAGE, TOTAL, STYLE, BENCH, TRACE, TWIST, STORM, STORE, STAKE, SMOKE, BLIND, UNITY, PAPER, TRIAL, SPOKE, GRADE, SWEAT, DRESS, PRINT, UNCLE, VERSE, SCREW, SHIRT, MIGHT, DREAM, PIECE, BROWN, BRUSH, FRUIT, CROSS, SPITE, VISIT, SUITE, REACH, RADAR, BAKER, AWARD, HOTEL, FEVER, BOARD, HOUSE, TOAST, ALERT, REBEL, CLAIM, SHEET, LEAVE

Appendix B

The quartile analyses presented below used the same analyses outlined in the main paper, unless otherwise indicated. The RTs were analysed using a one-way repeated-measures ANOVA, with quartile (first/second/third/fourth) as the factor. The analyses of RTs across all experiments were significant (all *F*'s > 30) and will not be discussed in detail; the data can be observed in Table B1. The "old" responses at test were analysed using a 4x2x2 repeated-measures ANOVA, with quartile (first/second/third/fourth) as one factor, repetition (repeated/non-repeated) as a second factor, and study-status (old/new) as the final factor. In all cases, only the highest-order interactions are reported for simplicity.

Table B1

Mean response times (ms) for each quartile within the test phase

	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Experiment 1	486 (162)	763 (213)	1072 (348)	2374 (1094)
Experiment 2	617 (205)	854 (322)	1188 (600)	2822 (1921)
Experiment 3	683 (95)	880 (151)	1214 (315)	2682 (1053)

Note: Table displays mean response times with standard deviation in parentheses

Experiment 1

The analysis of "old" responses (Figure B1) revealed a significant three-way interaction, $F(3,69) = 6.26$, $p < .001$, $\eta_p^2 = .546$. To understand this interaction, a 2x2 ANOVA was conducted for each quartile, with repetition and study-status as factors. The main effect of study-status was significant in all cases, all *p*'s < .001. None of the other effects were significant for quartile 1, all *p*'s > .10. For quartiles 2, 3, and 4, the interaction between repetition and study-status was significant, largest $p = .036$ ($\eta_p^2 = .177$). In all cases, the interaction was driven by a significant simple main effect for the old items, largest $p = .006$ ($d = .624$), indicating higher hits for non-repeated than repeated items. In all cases, the simple main effect for new items was not significant, smallest $p = .107$ ($d = .324$). Overall, the repetition decrement effect was not observed for the fastest recognition decisions, but was observed for recognition decisions made more slowly, which is broadly consistent with the quartile analysis for Experiment 4.

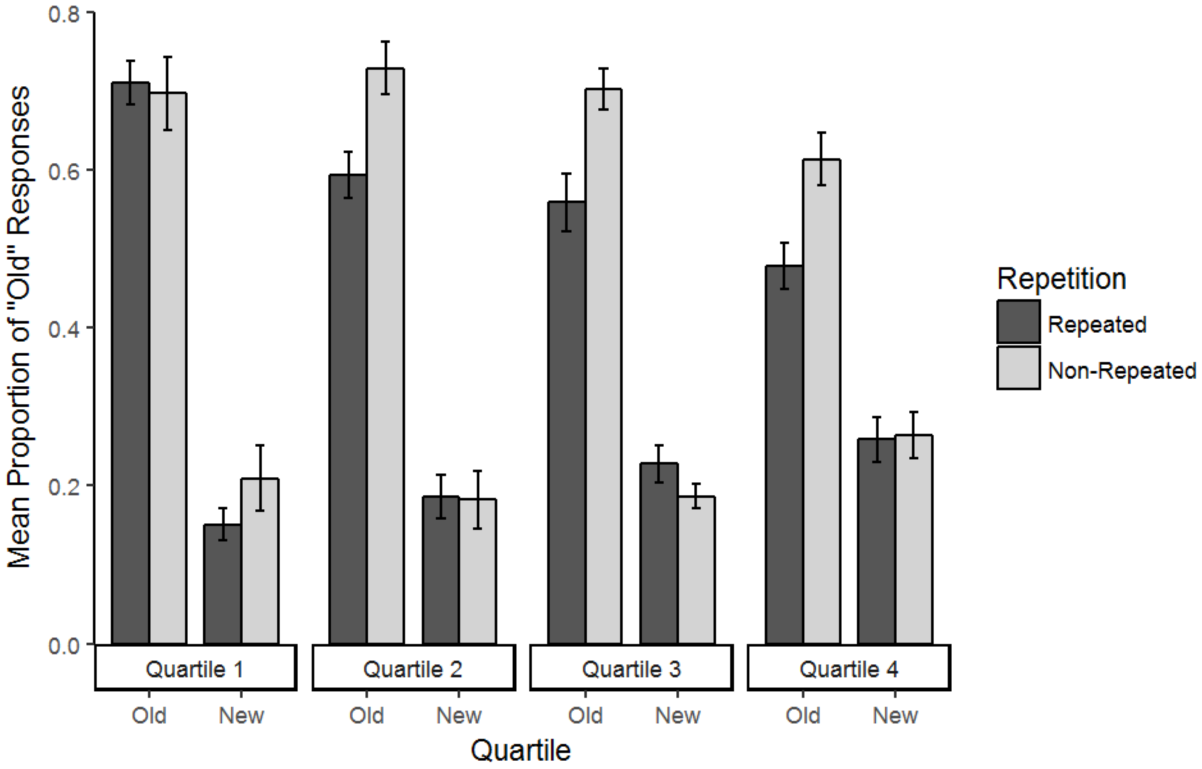

Figure B1: Quartile data for Experiment 1

Experiment 2

The analysis of “old” responses (Figure B2) revealed a significant interaction between quartile and repetition, $F(3,69) = 3.07$, $p = .033$, $\eta_p^2 = .317$. To understand this interaction, the effect of repetition was analysed separately for each quartile. In quartile 1, there were more “old” responses to repeated (.445) than non-repeated items (.347), $t(23)$

= 4.43, $p < .001$, $d = 0.904$. The corresponding effects for quartiles 2 and 3 were not significant, both t's < 1. Finally, the analysis for quartile 4 was significant, again revealing more "old" responses for repeated (.431) than non-repeated items (.383), $t(23) = 2.08$, $p = .048$, $d = .425$. Overall, the false recognition effect was present when responses were fast, and disappeared when responses slowed, though it did re-emerge during the slowest response quartile. Excluding the results of quartile 4, these results are largely consistent with those of Experiment 4, in which the false recognition effect disappears as responding slowed; however, no repetition decrement effect was observed in this experiment.

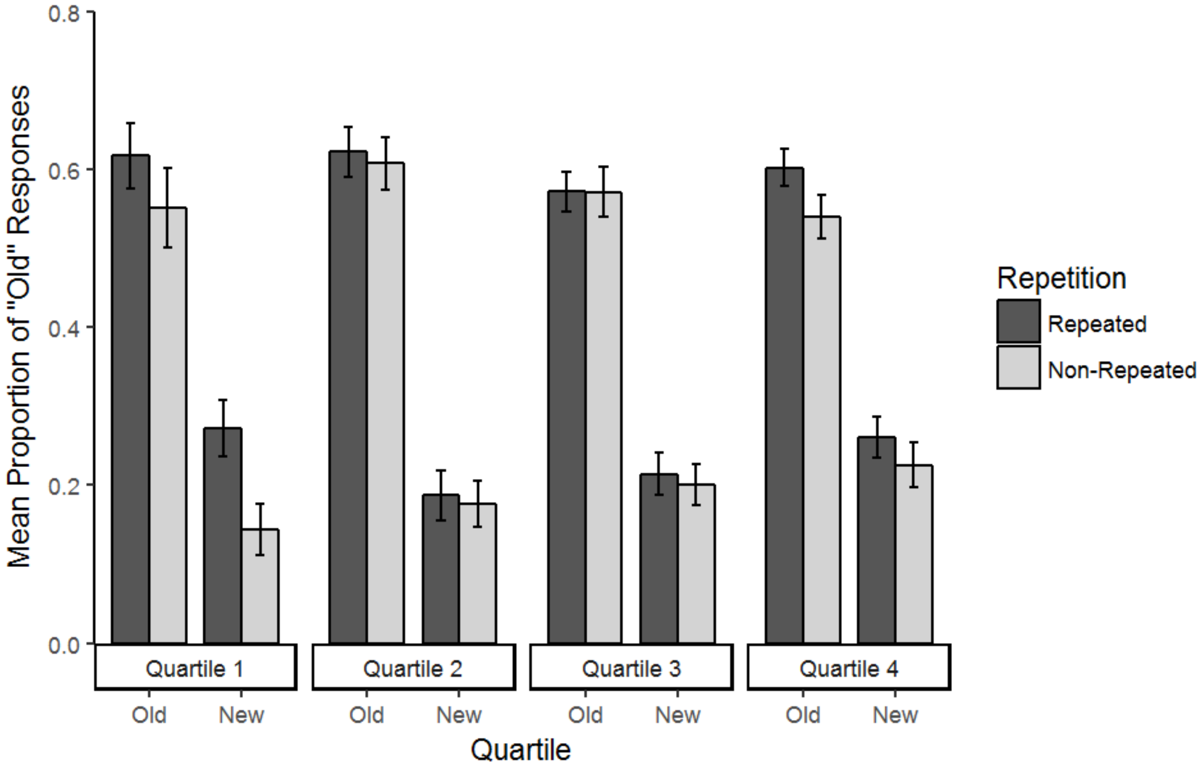

Figure B2: Quartile data for Experiment 2

Experiment 3

As repetition was not a factor at the time of test, only hit rates ("old" responses to old items) were analysed (Figure B3). As such, "old" responses were submitted to a 4x2 repeated-measures ANOVA, with quartile as one factor and repetition as another factor. There was a significant main effect of repetition, $F(1,47) = 10.75$, $p < .001$, $\eta_p^2 = .186$, indicating overall higher hits for non-repeated than repeated items. Although the interaction between quartile and repetition was not significant, $F(3,141) = 1.82$, $p = .148$, $\eta_p^2 = .134$, the repetition decrement effect appears to be negligible in quartile 1 and robust in quartiles 2 and 3. This trend is generally consistent with the results from Experiment 4, and the idea that the repetition decrement effect does not occur for the fastest recognition responses.

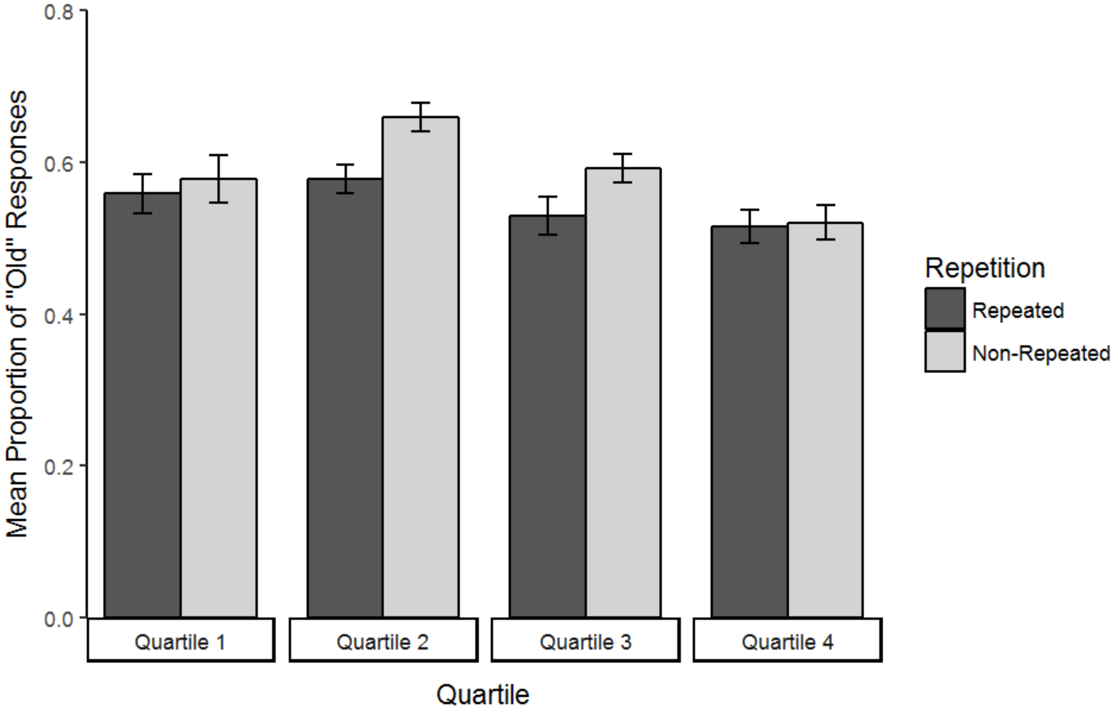

Figure B3: Quartile data for Experiment 3

Summary

Overall, the quartile analyses presented above are broadly consistent with the quartile analysis for Experiment 4. In experiments in which the repetition decrement effect was observed (Experiments 1 and 3), the repetition decrement effect emerged in later quartiles. In Experiments in which the false recognition effect was observed (Experiment 2), the effect is present for the fastest recognition decisions and decreases with increasing RTs. As a cautionary note, the results from quartile 4 are somewhat inconsistent across experiments.

CHAPTER 4: The role of expected fluency: A comparison of an old and a new false recognition effect

Rosner, T. M & Milliken, B. (Submitted).

Psychonomic Bulletin & Review

Manuscript ID: PBR-BR-18-124

Preface

Chapter 4 presents the results of two experiments in which two different false recognition effects were examined. The experiments conducted were primarily to better understand a false recognition effect in response to congruency that was originally reported by Rosner, D'Angelo, MacLellan, and Milliken (2015) who were interested in the impact of congruency on encoding. In Rosner et al. (2015), participants were presented with red and green interleaved words that were the same (congruent) or different (incongruent) at the time of study and test. They observed higher false alarms for congruent than incongruent items, an effect that was noted but not discussed as it was not the main focus of the investigation. The experiments in Chapter 4 aim to better understand this false recognition effect in response to congruency and determine if this effect is produced by similar processes that lead to the repetition-driven false recognition effect reported by Jacoby and Whitehouse (1989). While these effects *appear* similar, the results of these experiments suggest that each effect may be reflecting differences in how fluency is interpreted. It was observed that repetition-driven false recognition can be modulated by global proportion of item type, whereas this was not the case for congruency-driven false recognition. These experiments indicate that relative fluency of an item does not on its own determine how that fluency will be interpreted, but rather false recognition may be a product of many factors, one of which is relative fluency. While the motivation for this chapter was primarily driven by the goal to better understand the false recognition effect in response to congruency, the results obtained very much speak to the role of repetition in producing illusions of familiarity, and moreover, hold implications for the repetition decrement effect. These implications are discussed in further detail in Chapter 5 of this thesis.

Abstract

In a recent study examining how selective attention impacts recognition memory performance, a new method for measuring false recognition effects was serendipitously discovered (Rosner, D'Angelo, MacLellan & Milliken, 2015). This method involves interleaving words, with congruent items composed of the same red target and green distractor word and incongruent items composed of two different words. The present study was aimed at better understanding this false recognition effect, and in particular its relation to the false recognition effect introduced by Jacoby and Whitehouse (1989). In two experiments, the relative proportions of two item types was manipulated to produce different expectancies for processing fluency. Previous work has shown that expectation for processing fluency impacts the size of the Jacoby-Whitehouse effect, with a larger false recognition effect when repeated (fluent) items are more rare (Westerman, 2008). This result was replicated here with the brief-duration masked primes typically used to measure the Jacoby-Whitehouse effect. However, when congruent and incongruent items were used to manipulate fluency, the size of the false recognition effect was insensitive to the relative proportions of congruent and incongruent items. These findings highlight a new method for measuring false recognition, and demonstrate that this false recognition effect is controlled by different processes than control the Jacoby-Whitehouse false recognition effect.

Introduction

The feeling of recognizing someone, but not knowing *why*, is a ubiquitous experience. Passing an acquaintance on the street, watching a B-list actor in a movie, and encountering "the butcher on the bus" (Mandler, 1980) all hold the potential to produce this feeling-state. Much past research suggests that such feelings of familiarity have an attributional basis (Jacoby & Dallas, 1981). According to this view, fluent item processing is often attributed to that item having appeared in the past (Jacoby & Dallas, 1981).

However, processing fluency can be influenced by factors other than prior experience, such as stimulus clarity (Whittlesea, Jacoby, & Girard, 1990), semantic relatedness (Begg, Anas, & Farinacci, 1992), and conceptual expectancy (Whittlesea, 1993). The implication for a fluency attribution account is that feelings of familiarity can be produced "falsely" by factors other than prior experience (Jacoby, Kelley, & Dywan, 1989). The present study focuses on such false recognition effects. In particular, we introduce a new method for inducing false recognition, and contrast this new method with a well-studied and possibly related method (Jacoby & Whitehouse, 1989).

Processing Fluency and False Recognition

In a seminal study of false recognition, Jacoby and Whitehouse (1989) asked participants to learn a list of words for a later recognition test. During the recognition test, target words were preceded by prime words that were presented briefly and masked. The primes and targets were either the same word (repeated) or different words (non-repeated). The key result from this study was a higher false alarm rate for repeated than non-repeated target words. Jacoby and Whitehouse proposed that prime-target repetition increased processing fluency for targets, and that this fluency was misattributed to familiarity, driving up the false alarm rate for repeated relative to non-repeated items (Jacoby & Whitehouse, 1989). As the increase in false alarms is driven by attribution of fluency to an incorrect source, the effect is referred to as false recognition.

Whittlesea and Williams (1998, 2000, 2001) later pointed to the importance of relative rather than absolute processing fluency to feelings of familiarity. According to their proposal, feelings of familiarity occur when experienced processing fluency is discrepant from expected processing fluency. Westerman (2008) tested this idea using the Jacoby-Whitehouse method by manipulating the relative proportions of repeated and non-repeated items at the time of test. Participants in different groups experienced recognition test phases in which 10%, 33%, 67%, or 90% of the test items were repeated. The magnitude of the Jacoby-Whitehouse effect diminished linearly with increases in the proportion of repeated trials with the largest effect observed when repeated items were least frequent (Westerman, 2008). This result is consistent with the idea that feelings of familiarity are mediated by processing fluency relative to expectation rather than by absolute processing fluency.

The Present Study: A New Method For Measuring False Recognition

A recent study aimed at examining the link between selective attention and recognition uncovered a potential new method to study false recognition (Rosner, D’Angelo, et al., 2015). Rosner, D’Angelo et al. presented participants with red target words interleaved with green distractor words during a study phase and again during a following test phase. Examples of these stimuli can be seen in Figure 1. Half of the items were congruent (the red and green words were the same) and the other half were incongruent (the red and green words were different). The primary goal of this study was to examine the impact of conflict on memory encoding—recognition sensitivity was higher for incongruent than congruent items. Consequently, a second result from this study received little attention: false alarms were higher for congruent than incongruent items (see also Rosner, López-Benítez, D'Angelo, Thomson, & Milliken, 2018, supplementary material). The primary aim of the present study was to turn our attention to this effect, and in particular to examine its relation to the Jacoby-Whitehouse effect.

Figure 1. Examples of congruent (left) and incongruent (right) stimuli.

Experiment 1

The objective of this experiment was to determine whether processing fluency at test implemented with the congruency method of Rosner et al. (2015) produces a false recognition effect like that observed by Jacoby and Whitehouse (1989). To address this issue, a study list consisting of single words was presented to participants, which was followed by a recognition test phase that included intermixed congruent and incongruent items. If processing fluency for congruent and incongruent items at test can induce a false recognition effect, then the likelihood of an "old" response ought to be higher for congruent than incongruent items.

Assuming a false recognition effect would be observed, we were also interested in whether this effect could be attributed to the same processes as the Jacoby-Whitehouse false recognition effect. As noted above, Westerman (2008) demonstrated that the Jacoby-Whitehouse effect is larger when prime-target repetitions at test are rare than when they are common. Here, we examined whether the congruency-driven false recognition effect is also larger when congruent trials at test are rare than when they are common. To address this issue, two groups of participants were tested. For the low proportion congruent group, 20% of test items were congruent and 80% were incongruent. For the high proportion congruent group, 80% of test items were congruent and 20% were incongruent. If the false recognition effect elicited by congruency is produced by similar processes that elicit the Jacoby-Whitehouse effect, then a larger false recognition effect should be observed for the low proportion congruent group than for the high proportion congruent group.

Method

Participants. Thirty-two participants (22 female, mean age = 18.78) from the McMaster University student pool were recruited in exchange for course credit or $10. All participants had normal or corrected-to-normal visual acuity, normal colour vision, and spoke English fluently. Sixteen participants were randomly assigned to each of the high proportion congruent and low proportion congruent groups.

Stimulus and apparatus. The experiment was run on a Mac Mini computer and stimuli were displayed on a 24-inch BenQ LED monitor using PsychoPy software

(Peirce, 2007, 2009). Participants sat 50 cm from the monitor and were tested individually.

The stimuli consisted of 480 five-letter high-frequency nouns (Kučera & Francis, 1967). All stimuli were displayed in the centre of the screen against a black background. The study phase stimuli consisted of single white words with spaces between the letters to allow for interleaving of these same items in the test phase. Each word subtended a visual angle of 0.8º vertically and 5.9º horizontally. The test phase stimuli consisted of two interleaved words, one red and the other green (Figure 1). The two interleaved words together subtended visual angles of 1.0º vertically and 6.5º horizontally.

Design. The 480 words were randomly divided into eight lists of 60 words. These eight lists were further divided into two sets of four lists each. Each set was assigned to the role of target (red word at test) and distractor (green word at test) an equal number of times across participants. Within the target word set, roles were counterbalanced such that each list appeared as an old or new item an equal number of times. Within the distractor word set, lists were randomly assigned to the role of distractor for an old or new item. As we were primarily interested in examining false recognition effects, there were 60 old and 180 new items for each participant.

Words were randomly assigned to congruent or incongruent conditions with the following constraints. For the low proportion congruent group, 12 old items and 36 new items were congruent, and 48 old items and 144 new items were incongruent, resulting in 20% of the test items being congruent and 80% of the test items being incongruent. For the high proportion congruent group, 48 old items and 144 new items were congruent, and 12 old items and 36 new items were incongruent, resulting in 80% of test items being congruent and 20% of test items being incongruent. Congruent items consisted of identical red target and green distractor words, whereas incongruent items consisted of different red target and green distractor words. Incongruent distractor words were chosen randomly without replacement from the four lists making up the distractor word set. Finally, the relative placement of target and distractor words was counterbalanced, such that target words were presented on top for half of each item type.

Because incongruent items consist of two different words and congruent items consist of two versions of the same word, the number of total words presented to the low and high proportion congruent groups differed. Participants in the low proportion congruent group saw 432 unique words, whereas participants in the high proportion congruent group saw 288 unique words. Therefore, no participants saw the full set of 480 words, but each word was seen as a red target at test an equal number of times across participants.

Procedure. The experiment consisted of three phases: a study phase, a distractor math task, and a test phase. Participants were shown words one at a time in the study phase, and were asked to read aloud each word and remember it for a later memory test. Participants wore a headset with a microphone and their voices were recorded using Audacity software to ensure that all words were named. There were 60 trials in the study phase, each of which consisted of a fixation cross for 500 ms, a word for 2000 ms, and a blank screen for 250 ms.

Following the study phase, participants completed a distractor task consisting of basic arithmetic problems for five minutes. Arithmetic problems were presented on the computer display, and responses were written on scrap paper. Following the distractor task, participants completed the recognition memory test. A pair of interleaved words was presented on each trial, and participants were to discriminate whether or not the red word appeared earlier in the study phase. Each test phase trial began with a fixation cross for 500 ms. The target item then appeared and stayed on screen until response together with the words "OLD" and "NEW" on the bottom left and right of the screen, respectively, to remind participants of the response key mapping. Participants recorded their decision by pressing the A key for items judged "old" and the L key for items judged "new". There were 240 test phase trials, 60 of which were old and 180 of which were new.

Results

Proportions of items judged "old" in each condition are displayed in Figure 2. These data were submitted to a 2x2x2 mixed-factor ANOVA that treated congruency (congruent/incongruent) and study-status (old/new) as within-subject factors, and proportion congruent (low/high) as a between-subjects factor.

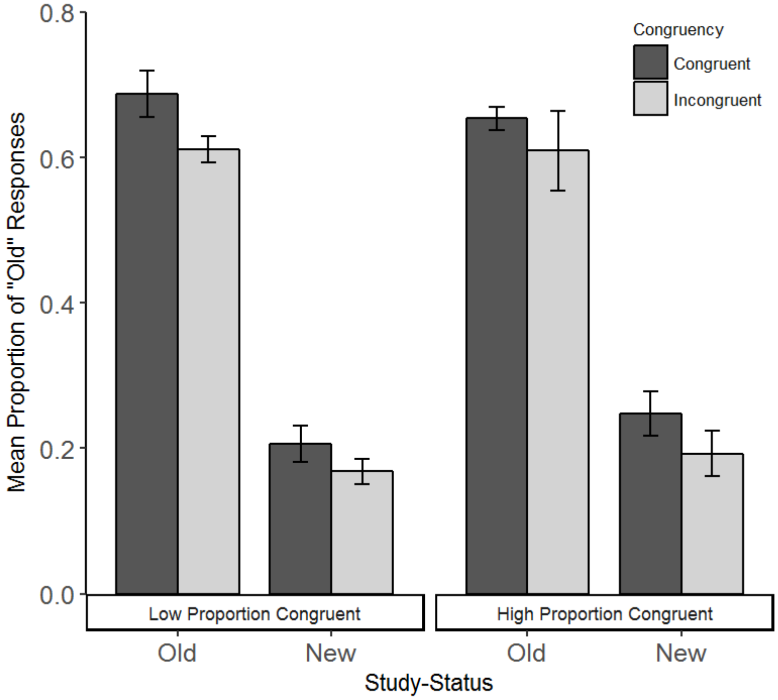

Figure 2. Mean proportion "old" for Experiment 1. Error bars here and on other graphs represent within-subject error corrected for between-subjects variability (Morey, 2008).

The main effect of study-status was significant, $F(1,30) = 221.30$, $p < .001$, $\eta_p^2 = .881$, with a greater proportion of "old" responses to old (.640) than new items (.204). More important, there was a significant main effect of congruency, $F(1,30) = 8.84$, $p = .006$, $\eta_p^2 = .228$, with more "old" responses to congruent (.449) than incongruent items (.395). Finally, and contrary to the hypothesis that proportion congruent would impact the size of the false recognition effect, the interaction between proportion congruent and congruency was not significant, $F < 1$. No other effects in the analysis were significant, all F's < 1.

As our primary interest was false recognition, false alarm rates were submitted to a separate 2x2 mixed-factor ANOVA that treated congruency and proportion congruent as factors. Consistent with the analysis reported above, only the main effect of congruency was significant, $F(1,30) = 10.77$, $p = .003$, $\eta_p^2 = .264$. The false alarm rate was higher for congruent (.227) than incongruent items (.180). Importantly, the interaction between proportion congruent and congruency was not significant, $F < 1$. This result indicates clearly that proportion congruent did not impact the false recognition effect.

Discussion

There were two important results in Experiment 1. First, the higher false alarm rate for congruent than incongruent items replicates the false recognition effect reported by Rosner et al. (2015). Second, this false recognition effect did not vary as a function of proportion congruent. If processing fluency relative to expectation drives this false recognition effect, as has been argued for the Jacoby-Whitehouse effect (Westerman, 2008), then the false recognition effect ought to have been larger for the low proportion congruent group than the high proportion congruent group.

Experiment 2

The results of Experiment 1 point to a difference between the congruency-based false recognition effect and the Jacoby-Whitehouse effect. Whereas the false recognition effect in Experiment 1 was not sensitive to the proportion of congruent items, the Jacoby-Whitehouse effect has been shown to be sensitive to the proportion of repeated items (Westerman, 2008). Both of these methods might reasonably affect expectation for processing fluency, and therefore it was surprising that these methods produced different results.

At the same time, the different results here and in the Westerman (2008) study could be due to idiosyncratic method differences between the studies. For example, Westerman tested 10%, 33%, 67%, and 90% repeated conditions, whereas we tested 20% and 80% congruent conditions. The narrower range of the expectancy manipulation in Experiment 1 relative to the study of Westerman (2008) could be responsible for the different results across the studies. Westerman (2008) also tested 60 participants per condition, whereas we tested 16 participants per condition. Lower power to detect an expectancy effect in Experiment 1 relative to the study of Westerman (2008) could also be responsible for the different results across the studies.

To address this issue, the goal of Experiment 2 was to examine whether the expectancy effect reported by Westerman (2008) would be observed with the parameters used in Experiment 1. To this end, Experiment 2 was identical to Experiment 1 with the exception that the masked priming method typically used to measure the Jacoby-Whitehouse effect replaced our congruency method. If proportion repeated influences the magnitude of the Jacoby-Whitehouse effect in this experiment, then we can conclude more confidently that the Jacoby-Whitehouse effect and our congruency false recognition effect are controlled by different processes.

Method

Participants. Thirty-two participants (26 female, mean age = 19.93) from the McMaster University student pool participated in exchange for course credit or $10. Sixteen participants were randomly assigned to each of the high proportion repeated and low proportion repeated conditions.

Stimuli, apparatus, design, and procedure. The stimuli, apparatus, design, and procedure were identical to Experiment 1 with the following exceptions. Words assigned to be distractors in Experiment 1 were instead assigned to be brief-duration masked primes in Experiment 2. Each test phase trial consisted of a 500 ms fixation cross, a 500 ms pre-mask (a row of five green X's), a 50 ms green prime word, a 500 ms post-mask (a row of five green X's), and a red target word, all presented in sequence at the center of the display. The target stayed on screen until a recognition response was provided.

Results

Proportions of items judged "old" are displayed in Figure 3. These data were submitted to a 2x2x2 mixed-factor ANOVA, with repetition (repeated/non-repeated) and study-status (old/new) as within-subject factors and proportion repeated (low/high) as a between-subjects factor.

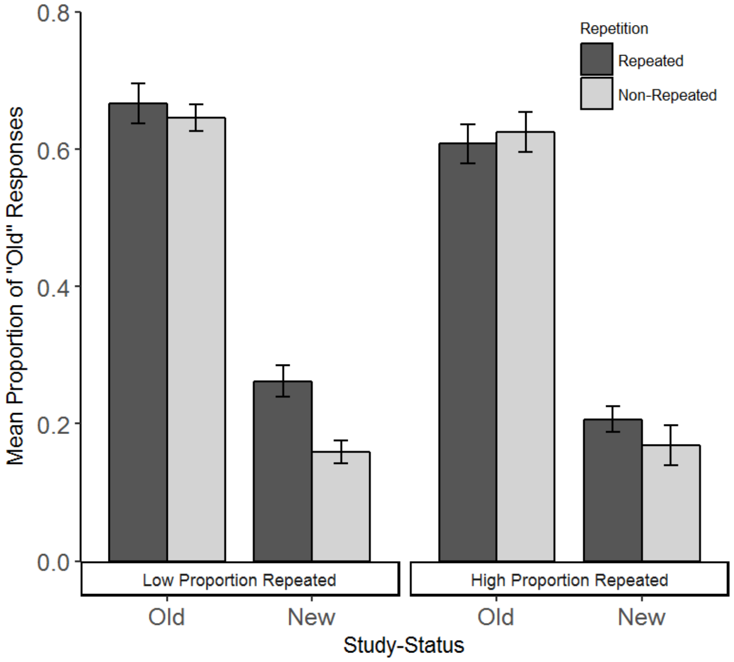

Figure 3. Mean proportion "old" responses for Experiment 2.

The main effect of study-status was significant, $F(1,30) = 373.35$, $p < .001$, $\eta_p^2 = .926$, with more “old” responses to old (.636) than new (.199) items. The main effect of repetition was significant, $F(1,30) = 5.95$, $p = .021$, $\eta_p^2 = .165$, but was qualified by a significant study-status by repetition interaction, $F(1,30) = 6.46$, $p = .016$, $\eta_p^2 = .177$. Separate analysis of the repetition effect for new items revealed a significant repetition effect, $t(31) = 4.38$, $p < .001$, $d = 0.75$, with a higher false alarm rate for repeated (.234) than non-repeated (.164) items. A corresponding analysis of old items was not significant, $t < 1$. The only other effect to approach significance was the interaction between proportion repeated and repetition, $F(1,30) = 2.98$, $p = .095$, $\eta_p^2 = .091$. The results in Figure 3 indicate a trend toward higher false alarm rates for the low proportion repeated group than for the high proportion repeated group. However, this key result is captured best by the following analysis of false alarm rates only.

To evaluate the false recognition effects of primary interest, false alarm rates were submitted to a separate 2x2 mixed factor ANOVA, with repetition and proportion repeated as factors. The main effect of repetition was significant, $F(1,30) = 21.37$, $p < .001$, $\eta_p^2 = .416$, with higher false alarms for repeated (.234) than non-repeated items (.164). Most important, there was a significant repetition by proportion repeated interaction, $F(1,30) = 4.52$, $p = .042$, $\eta_p^2 = .131$. For the low proportion repeated condition, the false alarm rate was higher for repeated (.262) than non-repeated items (.159), $t(15) = 4.20$, $p < .001$, $d = 1.05$. For the high proportion repeated condition, the false alarm rate was also higher for repeated (.206) than non-repeated items (.168), though this effect just approached significance, *t*(15) = 2.09, *p* = .054, *d* = .523. Together, these results indicate that the false recognition effect was larger for the low proportion repeated condition than the high proportion repeated condition.

Discussion

The results of Experiment 2 nicely replicate Westerman (2008). The Jacoby-Whitehouse false recognition effect was larger for the low proportion repeated condition than for the high proportion repeated condition. This result supports the idea that the Jacoby-Whitehouse effect is mediated by expected fluency, with larger effects when expected fluency is low (Whittlesea & Williams, 1998). Most important for our purpose, this replication of the Westerman (2008) study used parameters that were identical to those in Experiment 1. A comparison of the false recognition effects in Experiments 1 and 2 is presented in Figure 4. Together, the results of Experiments 1 and 2 suggest that processes mediating the congruency-based false recognition effect in Experiment 1 are not identical to those mediating the Jacoby-Whitehouse effect in Experiment 2.

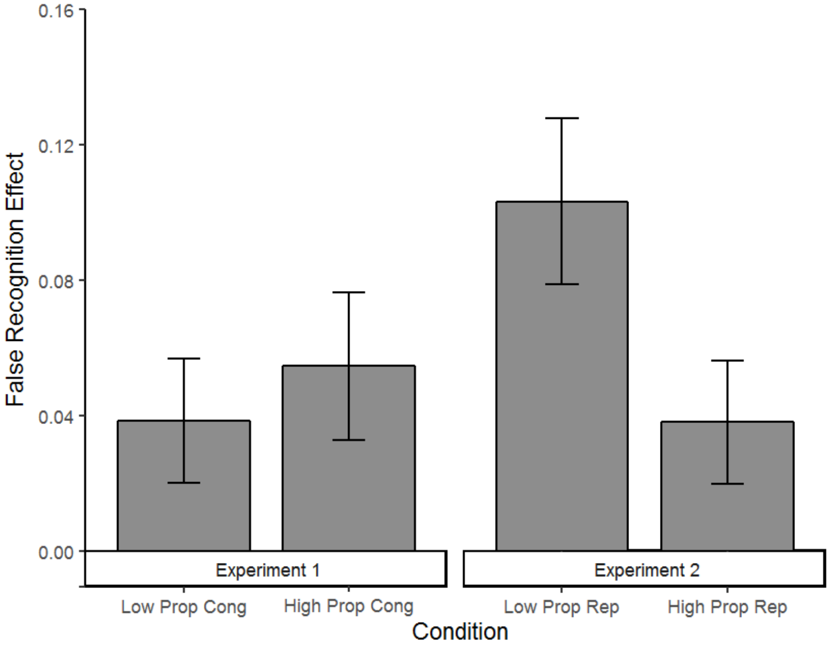

Figure 4. The false recognition effects (congruent – incongruent in Experiment 1; repeated – non-repeated in Experiment 2) in Experiments 1 and 2. Error bars reflect standard error of the mean.

General Discussion

The goal of the present study was to examine a new method for measuring false recognition. In Experiment 1, a congruency-based false recognition effect was observed using a procedure in which congruency was manipulated only at test. This false recognition effect did not vary as a function of the proportion of congruent items at test. Importantly, Westerman (2008) demonstrated that the Jacoby-Whitehouse effect varies inversely with the proportion of repeated items (Westerman, 2008), a result replicated in Experiment 2 using identical parameters to Experiment 1. Together, the results suggest that processes controlling the congruency-based illusion of familiarity are different from those controlling the Jacoby-Whitehouse effect.

Why does the global context set by the proportion of repeated items influence false recognition reliably, whereas the global context set by the proportion of congruent items shows no evidence of influencing false recognition? One possibility concerns the idiosyncratic nature of congruent/incongruent stimuli, and the idea that such unusual stimuli encourage item-specific rather than relational encoding processes (McDaniel & Bugg, 2008). If congruent and incongruent items are more unusual than single words, then congruent and incongruent items may be subject primarily to item-specific processing, whereas single words may also be subject to relational processing (Davis, Rosner, D'Angelo, MacLellan, & Milliken, 2018). Indeed, the fluency associated with repeated items stems from the relation between separate items (prime and target), whereas the fluency associated with congruency items stems from the structure of an item itself. With this in mind, the Jacoby-Whitehouse effect may be sensitive to the proportion of repeated items because of the opportunity to set an expectation based on the prime in relation to the following target. In contrast, the congruency-based false recognition effect may not be sensitive to the proportion of congruent items because each item is considered in isolation, with no opportunity to set an expectation in a way that would influence item processing. Future studies might aim to examine this distinction further. For example, based on the above conjecture, one would expect that false recognition effects based on stimulus clarity (e.g., Whittlesea, 1993) would not be sensitive to a proportion manipulation, as the fluency manipulation is item-based rather than relation-based.

In conclusion, this study highlights the idea that not all fluency-based false recognition effects are controlled by the same processes. Although both of the false recognition effects studied here may involve an attribution that resolves a feeling-state (Whittlesea & Williams, 1998), the details of processing that give rise to that feeling-state, and thus the broad attribution process, may differ depending on how processing fluency is implemented in the experiment. Going forward, it will be important to examine how manipulations of fluency differ from one another to better understand the attribution process that gives rise to false recognition effects, and by extension, the feeling-states in everyday life that we think of as remembering.

CHAPTER 5: General Discussion

There has been great interest in the field of cognition regarding the effect of stimulus repetition on encoding and memory performance. Prior work on the spacing effect has demonstrated that repetitions are better remembered when spaced apart than massed, suggesting that something about the processing of immediate repetitions undermines the benefit of repetition. Other researchers have been more interested in how repetition can produce an illusion of familiarity at test, showing that immediate repetition of items at retrieval increases the fluency of those items, creating the sensation that they have been recently experienced. Looking at these disparate literatures together hints at the idea that immediate repetition leading to increases in fluency (which is attributed to familiarity, as seen in the false recognition literature) may also be an explanatory mechanism for the spacing effect in which more fluent massed items are not encoded as well as spaced items.

Prior to the work reported in this dissertation, the relation between repetition at study and repetition at test had not been formally addressed, with these two bodies of work staying decidedly separate. This separation is understandable. Those studying the spacing effect would likely not want to manipulate conditions at test, for doing so may cloud the influence of repetition at the time of encoding. Alternatively, those interested in false recognition effects by definition are mostly interested in the influence of various manipulations at test, and as such, those manipulations during encoding would simply get in the way of measuring the desired effect. However, understanding the effects of repetition at study and test in tandem may help us better understand how these effects are related and can provide insight into human cognition as a whole.

The goal of the current thesis was first to better understand the role of repetition during encoding. When this research programme began, no work had been done examining how immediate repetition, as compared to a single encounter with a stimulus, impacts memory. This comparison aimed at better understanding processing that determines the extent to which resources are dedicated to encoding, and perhaps also to better understanding of the spacing effect. The second goal of this thesis was to look at effects of repetition on encoding and retrieval in tandem in order to shed light on similar processes that may be leading to seemingly separate effects.

In Chapter 2, initial experiments regarding stimulus repetition on remembering were reported. Across three experiments, a novel finding was observed: better memory for non-repeated than repeated items. This *repetition decrement effect* was only disrupted in Experiment 4, in which participants were asked to read both prime and target words. In this experiment, there was better memory for repeated than non-repeated words, but the size of the memory benefit was small. Taken together, the findings in this chapter demonstrate that immediate repetition is detrimental to memory, particularly when prime words are ignored (see Collins, Rosner, & Milliken, 2018 for more on prime encoding demands).

The method used in Chapter 2 pointed toward a systematic investigation of the effects of repetition at study and test, an issue examined in Chapter 3. In this chapter, I aimed to demonstrate that the repetition decrement effect and the repetition-based false recognition effect are due to similar processing. Namely, it was hypothesized that fluency indicates that an item is old; when driven by repetition, this cue at study results in poor encoding (repetition decrement effect) and at test results in illusions of familiarity (false recognition effect). After exploring the conditions that produce each effect, it was suggested that the expression of the repetition decrement effect is dependent on the presence (or lack thereof) of a false recognition effect, indicating a possible time course for each effect. This hypothesis was explored in a quartile analysis of Experiment 4 and confirmed using a response-deadline procedure in Experiment 5, demonstrating that when fluency-based cues could be used at test, faster responding results in the false recognition effect and slower responding results in the repetition decrement effect. These results demonstrate that although similar cues may lead to the two effects, which effect is observed is dependent on response strategy at test. If responses are made quickly, they will be driven by fluency cues at test, resulting in a false recognition effect. If responses at test are made slowly and effortful retrieval can occur, the impact of fluency on encoding will be observed.

Chapter 4 explored false recognition effects more closely, with the specific goal of contrasting a repetition-based false recognition effect with a congruency-based false recognition effect. The results of two experiments demonstrate that although these two effects may be due to fluency attributions, the importance of relative fluency to finding a false recognition effect may differ depending on the stimuli being used. Although the main focus of this chapter was to better understand the congruency-based false recognition effect, the results demonstrate how relative fluency is important in observing a repetition-based false recognition effect. More important, these results hold implications for the repetition decrement effect, suggesting that relative fluency may play a role during encoding, a possibility explored below.

It should be noted that some of the results reported in Chapters 2 and 3 may be seen as being at odds with one another. Specifically, in Experiments 1 and 2 of Chapter 2 in which a remember/know procedure was employed, I report that the repetition decrement effect appears to be driven by familiarity and not recollection. However, the results from Experiment 5 of Chapter 3 in which a response deadline procedure is used shows that the repetition decrement effect appears when responding at test is slower, which is indicative of recollection rather than familiarity (e.g., Jacoby & Dallas, 1981; Quamme et al., 2007). Taken together, these results appear to be in opposition with one another. However, the difference in familiarity estimates reported in Chapter 2 has not been replicated in other work. In fact, using a similar procedure, Collins et al. (2018) reported that the repetition decrement effect appeared to be driven by differences in recollection, rather than familiarity. Therefore, the results from Chapter 2 that contradict the findings in Chapter 3 may not be reliable. Future work should aim to better understand the retrieval processes that contribute to the repetition decrement effect to clarify these disparate findings.

The results of the empirical chapters in this thesis all explore the role of fluency and how fluent information is processed. This work points to a connection between the repetition decrement and false recognition effects, with fluency being used as a cue in a similar way both when encoding information and in creating feelings of familiarity. Next, I will turn to a more detailed discussion regarding this connection between the effects in question and the support for this framework, as well as possible alternate accounts for the reported findings.

Fluency as a Cue

As stated, the work in this thesis hints at the use of fluency as a cue for "oldness" (or disfluency as a cue for "newness"). It was suggested that fluency may be a signal that indicates whether or not information is already known. If fluency is experienced (indicating the item is known), then the item will not be encoded. Alternatively, disfluency may be a cue that an item needs to be better learned, resulting in increased encoding effort. This idea is well-demonstrated by the repetition decrement effect: better memory for non-repeated than repeated items, presumably due to the fluent processing of repetitions (Jacoby & Dallas, 1981). Of course, when this same fluency signal is experienced at test, an attribution is made regarding the item in question and there is a greater probability of an "old" response, producing a false recognition effect (Jacoby & Whitehouse, 1989). I have also suggested that this cue at test that signals "oldness" or "newness" triggers the same encoding processes evident during the study phase. In other words, as mentioned in Chapter 3, this framework would predict that non-repeated items at test would be better encoded than repeated items at test. This hypothesis could easily be evaluated by including a second test phase after the initial test phase. In this second recognition test phase, participants' could be probed with repeated and non-repeated targets from the first test phase. Observing the repetition decrement effect in this context would support the notion that fluency can be used as a cue to encode and can also be used to make attributions about various aspects of an item (such as whether or not it has been experienced previously).

However, this study has yet to be conducted, and as such, support for the framework outlined in this thesis is less direct. In Chapter 3, both the repetition decrement effect and the false recognition effect were observed in similar contexts and with similar manipulations, and thus an assumption of similar mechanisms at both stages is a reasonable one. It could, of course, be possible that the two effects being studied are due to entirely different processes, both of which happen to arise when repetition is manipulated (indeed, it may even be the case that fluency is not the driving force behind both effects). Although the empirical evidence presented does not fully rule out this latter view, it does seem improbable given that the repetition decrement effect has only been studied using an incidental encoding paradigm. Since participants were unaware of an upcoming memory test, any differences in memory are likely due to automatic encoding processes, rather than being driven by intentional encoding effort. Therefore, any processes that led to a difference in memory for repeated and non-repeated items could not be because of specific intent to encode, but rather because these processes are *always* operational. If one agrees with this assumption, it should therefore hold that these same processes would be operational during the test phase as well. In sum, although support for the framework outlined here is indirect, it is still reasonable to conclude that the repetition decrement and false recognition effects are a result of the same fluency-based cue. Further work should be done to more directly test this hypothesis, as doing so would reveal more about the basic mechanisms that drive our encoding of information and behaviour on a regular basis.

Furthermore, this framework leads to specific predictions in light of the results of Chapter 4, in which the repetition-based false recognition effect was attenuated when repetitions were common. According to Westerman (2008), this attenuation of the false recognition effect is due to reduced relative fluency for repeated items; since repetitions (and fluency) are common, there is a reduced impact of fluency on behaviour, possibly because expected fluency matches actual fluency (Whittlesea & Williams, 1998, 2000, 2001). Connecting these results to the repetition decrement effect, it then follows that as relative fluency decreases (i.e., repetitions become more common), then the repetition decrement effect may also be attenuated. Fluency becoming expected in a given situation may indicate it is not a valid cue for an item being known, and disfluency would no longer be used as a cue for encoding. Therefore, a smaller repetition decrement effect should be observed. Future studies should aim to examine the impact of proportion of repeated items at study on the repetition decrement effect. Observing a smaller repetition decrement effect when repetitions are common than when repetitions are rare would further support the notion that the repetition decrement and false recognition effects are related.

Of course, this framework has implications for other effects of repetition reported in the literature. Next, I aim to discuss how the findings presented in this thesis may shed light on the processes that result in the spacing effect and explore in more detail how the spacing and repetition decrement effects may be related. Then, I will turn to an exploration of the relationship between the repetition decrement effect and the repetition suppression effect, a finding of reduced neural activity in response to repetitions.

The Repetition Decrement Effect and the Spacing Effect

The framework outlined in this dissertation may speak to the mechanisms driving the spacing effect. The possibility that fluency itself leads to poor encoding can help account for the finding of better memory for spaced than massed repetitions. It is clear that massed repetitions demonstrate the detriment of immediate repetition to memory, especially when considered in the context of the memory strength effect (better memory for items seen twice than items seen once when spaced). That is, we can conclude that seeing an item twice over time is better than seeing an item once, and that massing repetitions is not as effective as spacing them.

It may be that massed repetitions are not well-remembered because repeated items are processed fluently, and thus are subject to diminished encoding. In effect, massed repetitions in spacing effect studies may produce memory performance that is equivalent to a single exposure, as little more than the first encounter with an item may actually be remembered. On the other hand, spaced repetitions may be better remembered because repeated items are not as fluently processed, and are therefore not subject to diminished encoding. This conclusion is supported by the results of Experiment 4 in Chapter 2. Recall that in this experiment, two groups of participants encountered study phase trials each containing two green primes and a red target. The lag-0 group experienced massed repetitions (for half the trials, the second prime and target were the same) and the lag-1 group experienced spaced repetitions (for half the trials, the first prime and the target were the same). Importantly, participants read all three words aloud, meaning they were attending to primes and targets equally. This instructional manipulation allows this experiment to be a better proxy for spacing effect studies, in which all words are attended to throughout the study phase, and led to the elimination of the repetition decrement effect. Specifically, the effect was reversed for the lag-1 group, with better memory for spaced repetitions than non-repetitions. However, for the lag-0 group, the difference in memory between repeated and non-repeated items was not significant: memory was equivalent for massed repetitions and non-repetitions. Similar findings were reported by Collins et al. (2018). These results demonstrate that massed repetitions may not provide a memory benefit above and beyond a single exposure. Moreover, these findings suggest that the spacing effect may be partially due to the fluency afforded to the second instance of a massed repetition resulting in poor encoding of the repetition.

Of course, most of the findings reported in this thesis do not apply directly to the spacing effect. After all, in spacing effect studies, participants typically attend to all words equally, whereas in most of the experiments reported here, participants ignored the prime words. Moreover, observing the repetition decrement effect is highly dependent on whether or not prime words are attended to. This finding is more concretely demonstrated in experiments in which the effects of prime encoding demands are directly manipulated, showing that paying greater attention to prime words results in the elimination or even reversal of the repetition decrement effect (Collins et al., 2018). However, more recent work suggests that even when prime words are ignored, a benefit of spacing repetitions can be observed (Collins & Milliken, 2018). In the key reported experiment, participants completed two study phases followed by one recognition memory test. The first study phase was identical to most of the experiments reported in this thesis: participants saw a green prime followed by a red target, and were asked to name the red target aloud. Half of the words were repeated (i.e., massed repetitions) and half of the words were non-repeated. This first phase was designed to compare the effect of immediate repetitions and immediate alternations on recognition memory, and should produce the repetition decrement effect. In the second study phase, which occurred approximately 10 minutes after the first study phase, participants were simply presented with red target words which they named aloud. Critically, this second study phase also had two types of trials. Non-repeated trials were words that had not yet been presented in the experiment. Repeated trials were words that had been shown as green primes for non-repeated trials in the first study phase (i.e., spaced repetitions). Finally, participants were given a surprise recognition memory test.

This clever design made possible a comparison of memory for spaced repetitions (repeated in the second phase) with massed repetitions (repeated in the first phase), as well as a comparison of repeated and non-repeated trials for both immediate and spaced repetitions. Collins and Milliken (2018) replicated the repetition decrement effect, with better recognition sensitivity for non-repeated targets than repeated targets from the first phase. The repetition decrement effect was reversed for items in the second study phase, with better recognition sensitivity for repeated targets than non-repeated targets. Most interesting was the comparison of items repeated in the first study phase (massed repetitions) and items repeated in the second study phase (spaced repetitions). Despite the first instance being ignored for both types of items, Collins and Milliken (2018) observed the classic spacing effect: better recognition sensitivity for spaced repetitions than massed repetitions. These results confirm that repetition is detrimental to memory when repetitions are massed but beneficial to memory when they are spaced. Moreover, it demonstrates a clear connection between massed items in the spacing effect literature and repeated items in the repetition decrement effect, suggesting that the framework outlined here may be a suitable one for the spacing effect.

To summarize, the findings of the experiments reported in this thesis (see also Collins & Milliken, 2018; Collins et al., 2018) point to processes that may drive the well-studied spacing effect. Increased fluency for the second instance of a massed repetition may lead to a reduced benefit of repetition when compared to spaced items. In effect, massed repetitions may produce equivalent recognition performance to a single instance in the spacing effect literature. This notion is supported in Chapter 2 of this thesis (and by Collins et al., 2018), and can even be observed when examining the effect of ignored primes on both immediate and spaced repetitions (Collins & Milliken, 2018).

The Repetition Decrement Effect and the Repetition Suppression Effect

Another well-studied effect in which immediate repetitions are compared to immediate alternations is the repetition suppression effect. This is the finding that neural activity is reduced when stimuli are repeated (see Gotts, 2016; Henson & Rugg, 2003 for reviews), with the main goal of these studies being to better understand the neural underpinnings of priming effects in behaviour. Although it has yet to be examined empirically, studies of the repetition suppression effect may point to a neural basis for the repetition decrement effect (see also Henson & Gagnepain, 2010). It could be that reduced neural activity for repetitions is responsible for the repetition decrement effect, or it could be that poor encoding of repetitions results in reduced neural activity for repetitions.

Interestingly, there is an established connection between the repetition suppression and spacing effects, with prior work showing greater repetition suppression (i.e., reduced neural activity) for massed than spaced repetitions (Xue et al., 2011; see also Zhao et al., 2015). Moreover, Xue et al. (2011) demonstrated that for both spaced and massed repetitions, activity was reduced (that is, repetition suppression was greater) for forgotten than remembered items. These results point to a neural basis of the spacing effect, demonstrating that repetition suppression is related to poor memory later on. Extending these results to the repetition decrement effect, it is reasonable to predict that greater repetition suppression would be observed for repeated than non-repeated items, and more interestingly, that the level of suppression might predict remembered and forgotten items. Future work should aim to examine this possible connection between the two effects to establish if differences in repetition suppression can be observed using the paradigm described here. If a link between the repetition decrement and repetition suppression effects can be observed, then more work should be done to determine the causal relation between these effects. Does reduced encoding of repetitions lead to neural suppression, or is neural suppression the driving factor for the behavioural effect observed in memory? It may also be the case that the two effects are correlated, with a third unconsidered factor accounting for each of these effects. Regardless of the causal direction, observing a connection between the repetition decrement and repetition suppression effects would not only provide a possible neural basis for the impact of repetition on memory, but also further confirm the connection between the repetition decrement and spacing effects.

The Impact of Proportion Repeated Items

Interestingly, research examining the repetition suppression effect shows that the amount of suppression observed can be attenuated by adjusting the proportion of repetitions and alternations (Summerfield, Trittschuh, Monti, Mesulam, & Egner, 2008). In this experiment, participants experienced blocks of trials in which repetitions were common (75% repeated) or rare (25% repeated). Summerfield et al. (2008) observed repetition suppression in the 75% repeated block but not in the 25% repeated block (see also Larsson & Smith, 2012; Summerfield, Wyart, Johnen, & Gardelle, 2011). That is, when repetitions were common, there was less neural activity for repetitions than for alternations; this difference was not observed when repetitions were rare. This result has been used to suggest that repetition suppression is not just a result of bottom-up processes, and can be adjusted based on expectation (Summerfield et al., 2008).

More important, the research by Summerfield et al. (2008) suggests one of two possibilities regarding the relation between the repetition decrement effect and the repetition suppression effect. Either (1) the two effects are not related or (2) the two effects are related, but the framework outlined in this dissertation connecting the repetition decrement effect and the false recognition effect may be incorrect. These conclusions are drawn by examining the experiments presented in Chapter 4 in the context of the work reported by Summerfield et al. (2008). Recall that in Chapter 4, a repetition-based false recognition effect was attenuated when there was a higher proportion of repeated than non-repeated items at test. As suggested previously, this finding may be extended to the repetition decrement effect, with the prediction that the effect will become smaller (i.e., memory performance will be equivalent for repeated and non-repeated items) as the proportion of repetitions is increased. If the repetition decrement and repetition suppression effects are related, one might also predict reduced repetition suppression as repetitions become more common, as this would align with the predictions for memory performance. Of course, this prediction is in opposition to the results reported in the literature, in which repetition suppression is *greater* when repetitions are common than when repetitions are rare (Larsson & Smith, 2012; Summerfield et al., 2008, 2011).

Therefore, (at least) one of these predictions is likely to be incorrect. The repetition suppression and the repetition decrement effects may be unrelated, and as repetitions become more common, repetition suppression may increase (difference in neural activity between repetitions and non-repetitions becomes larger) while the repetition decrement effect is attenuated (difference in memory between repetitions and non-repetitions becomes smaller). Alternatively, it may be that the repetition suppression and repetition decrement effects are related, with both increasing in size as repetitions become more common. However, this finding would indicate that the repetition decrement and false recognition effects are unrelated, as the false recognition effect does appear to reduce in size as repetitions become more common.

Of course, more work must be done before any of these conclusions can be drawn with confidence. For now, there appears to be a connection between the repetition decrement and repetition suppression effects. This connection must be explored in more detail to better understand the implications for the repetition decrement effect and how it may or may not relate to effects of false memory and repetition suppression.

Future Directions and Open Issues

Though some suggestions for future work have already been provided, there are still many other avenues to explore regarding the impact of immediate repetition on memory. Given that this thesis provides some of the only examples of the repetition decrement effect (see also Collins & Milliken, 2018; Collins et al., 2018), there is still much that is unknown about this effect. One issue yet to be addressed is whether or not this effect would be produced under intentional learning situations. If participants are aware of an upcoming memory test, would they be able to use the repetition of items at study to their advantage when encoding? Moreover, would better memory for repeated than non-repeated items in these circumstances reflect an impact of encoding intentionality, or would it simply be due to more attention paid to primes (e.g., Collins et al., 2018) as they provide useful information 50% of the time? The answers to these questions are relevant to an understanding of the automatic basis of encoding and the cues that we use both intentionally and unintentionally when learning. Moreover, research on intentionality may one day be extended to practical applications. If the repetition decrement effect is still observed when learning is intentional, then it would be useful to see how this effect generalizes to an educational setting. Is immediate repetition also detrimental when learning more complex information that may be encountered in a classroom? If so, how can we use the principles pointed to by the repetition decrement effect to improve learning and information retention? For now, there is still much to be explored regarding the repetition decrement effect, though this research may one day help inform teaching and learning strategies.

It would also be useful to learn if the repetition decrement effect can be observed when employing a recall test rather than a recognition test. Prior work has suggested that the direction of various memory effects is highly dependent on whether a recognition or recall test is used. A classic example is the word frequency effect (Gregg, 1976; see also Gillund & Shiffrin, 1984; Glanzer & Adams, 1985; Higham, Perfect, & Bruno, 2009; Joordens & Hockley, 2000; Kinoshita, 1995; Kinsbourne & George, 1974). In studies of the word frequency effect, people are given high and low frequency words to read and later remember. On recognition memory tests, people show better memory for low than high frequency words. On free recall tests, people often produce the opposite pattern of results: better memory for high than low frequency words (particularly if the words are presented in a pure list at study, see McDaniel and Bugg, 2008 for a review). Similar results are found with effects of perceptual interference (better memory for masked items than intact items on a test of recognition and the opposite for pure lists in free recall; Hirshman & Mulligan, 1991; Mulligan, 1999; Nairne, 1988); and the generation effect (better memory for generated than read items on a test of recognition and the opposite for pure lists in free recall; Nairne, Riegler, & Serra, 1991; Slamecka & Graf, 1978). One proposed framework to reconcile all of these effects suggests that difficult processing increases item-specific learning at the cost of relational learning (how individual items relate to each other; McDaniel & Bugg, 2008). On a recognition test, item-specific learning can be quite beneficial; when a target is provided, only information about the item itself can support retrieval, and relational information will be less useful. As such, words that trigger item-specific learning at study (e.g., low frequency words, generated words, and masked words) ought to be better remembered on a recognition test than words that trigger relational learning at study (e.g., high frequency words and intact words). In a test of free recall, however, relational learning would provide a large benefit. Since items must be self-generated in a recall task, items that were connected to other items at study would be more likely to be generated in the first place. As such, items that encourage relational learning at study (particularly in a pure-list format; McDaniel & Bugg, 2008) ought to be recalled with greater likelihood, as stronger connections are formed between items, which in turn facilitates the generation of candidate items to be recalled at test.

Turning back to the repetition decrement effect, similar principles may apply. For example, non-repeated items may trigger item-specific learning whereas repeated items could favour relational encoding. If this is the case, then it is not surprising that the repetition decrement effect is observed in recognition memory tasks, which tap primarily item-specific learning. Using a free recall task instead would be useful to better understand the impact of repetition on memory, particularly if using a pure-list format at study. If more repeated than non-repeated items are recalled, then the repetition decrement effect could easily fit into the framework provided by McDaniel and Bugg

(2008). If the classic repetition decrement effect is still observed, however (i.e., better recall of non-repeated than repeated items), then it would point to repetitions leading to different encoding processes than those that have been outlined for other memory effects. Regardless of what is found, the suggestion presented in this dissertation that the repetition decrement effect occurs in response to a fluency cue could fit into the framework provided by McDaniel and Bugg (2008). Better memory for repeated than non-repeated items in a recall task would indicate that disfluency is a signal for item-specific encoding in particular, rather than both item-specific and relational information. Alternatively, better memory for non-repeated than repeated items in a recall task would indicate that disfluency is a signal for both types of learning, in which case the framework offered to explain the repetition decrement effect would not necessarily explain other effects in the literature. Again, knowing the answer to all of these questions can help to identify how the repetition decrement is related to other well-studied memory effects.

Conclusion

In summary, this thesis describes the first demonstration of the repetition decrement effect: better memory for items shown once (non-repetitions) than items shown twice (repetitions) in immediate succession. Moreover, a framework has been proposed to connect this finding to classic effects of false recognition. According to this framework, both the repetition decrement effect and the repetition-based false recognition effect are produced in response to fluency. Specifically, this framework is borne out of the finding that information that is already known is processed more fluently than new information, meaning that fluency may be used as a cue for “oldness” (Jacoby & Dallas, 1981). Therefore, the repetition decrement effect may occur due to increased fluency for repeated items at study signalling that those items are already known, resulting in poor encoding for those items compared to disfluent non-repeated items. If repetition is manipulated at test, that same signal in response to fluency for repeated items results in increased “old” responses for those items compared to non-repeated items. These findings may also be applied to the classic spacing effect (better memory for spaced repetitions than massed repetitions) and may be connected to the repetition suppression effect (reduced neural activity for repetitions than alternations). Future research should aim to better understand the mechanisms behind the repetition decrement effect, and further explore how this effect is connected to other effects of repetition and memory in the literature. For now, this research demonstrates that more exposure is not always beneficial and contradicts the intuitive notion that repetition necessarily leads to better memory.

References

Alter, A. L., & Oppenheimer, D. M. (2006). Predicting short-term stock fluctuations by using processing fluency. *Proceedings of the National Academy of Sciences*, *103*(24), 9369–9372. http://doi.org/10.1073/pnas.0601071103

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review*, *13*(3), 219–235. http://doi.org/10.1177/1088868309341564

Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming Intuition: Metacognitive Difficulty Activates Analytic Reasoning. *Journal of Experimental Psychology: General*, *136*(4), 569–576. http://doi.org/10.1037/0096-3445.136.4.569

Baddeley, A. D., & Longman, D. J. A. (1978). The influence of length and frequency of training session on the rate of learning to type. *Ergonomics*, *21*(8), 627–635. http://doi.org/10.1080/00140137808931764

Bahrick, H. P. (1979). Maintenance of knowledge: Questions about memory we forgot to ask. *Journal of Experimental Psychology: General*, *108*(3), 296–308. http://doi.org/10.1037//0096-3445.108.3.296

Bahrick, H. P., & Phelps, E. (1987). Retention of Spanish Vocabulary Over 8 Years. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(2), 344–349.

Banks, W. (1970). Signal Detection Theory and human memory. *Psychological Bulletin*, *74*(2), 81–99. http://doi.org/10.1037/h0029531

Begg, I. M., Anas, A., & Farinacci, S. (1992). Dissociation of Processes in Belief: Source Recollection, Statement Familiarity, and the Illusion of Truth. *Journal of Experimental Psychology: General*, *121*(4), 446–458. http://doi.org/10.1037/0096-3445.121.4.446

Bernstein, I. H., & Welch, K. R. (1991). Awareness, false recognition, and the Jacoby-Whitehouse effect. *Journal of Experimental Psychology: General*, *120*(3), 324–328. http://doi.org/http://dx.doi.org/10.1037/0096-3445.120.3.324

Bjork, E. L., & Bjork, R. A. (2011). Making Things Hard on Yourself, But in a Good Way: Creating Desirable Difficulties to Enhance Learning. In *Psychology and the Real World: Essays Illustrating Fundamental Contributions to Society* (pp. 55–64). Worth Publishers.

Bjork, R. A. (1994). Memory and Metamemory Considerations in the Training of Human Beings. In J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about Knowing* (pp. 185–205). Cambridge, MA: MIT Press.

Bjork, R. A., & Allen, T. W. (1970). The Spacing Effect: Consolidation or Differential Encoding? *Journal of Verbal Learning and Verbal Behavior*, *9*(5), 567–572. http://doi.org/10.1016/S0022-5371(70)80103-7

Boldini, A., Russo, R., & Avons, S. E. (2004). One process is not enough! A speed-accuracy tradeoff study of recognition memory. *Psychonomic Bulletin & Review*, *11*(2), 353–361. http://doi.org/10.3758/BF03196582

Boldini, A., Russo, R., Punia, S., & Avons, S. E. (2007). Reversing the picture superiority effect: a speed-accuracy trade-off study of recognition memory. *Memory*

& Cognition, *35*(1), 113–23. http://doi.org/10.3758/BF03195948

Braun, K., & Rubin, D. C. (1998). The spacing effect depends on an encoding deficit, retrieval, and time in working memory: evidence from once-presented words. *Memory*, *6*(1), 37–65. http://doi.org/10.1080/741941599

Bruno, D., Higham, P. A., & Perfect, T. J. (2009). Global subjective memorability and the strength-based mirror effect in recognition memory. *Memory and Cognition*, *37*(6), 807–818. http://doi.org/10.3758/MC.37.6.807

Collins, R. N., & Milliken, B. (2018). The Repetition Decrement and Spacing Effects: Deficient Processing of Primed Words. *Manuscript Submitted for Publication*.

Collins, R. N., Rosner, T. M., & Milliken, B. (2018). Remembering “primed” words: The effect of prime encoding demands. *Canadian Journal of Experimental Psychology*, *72*(1), 9–23. http://doi.org/10.1037/CEP0000138

Craik, F. I. M., Govoni, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General*, *125*(2), 159–180. http://doi.org/10.1037/0096-3445.125.2.159

Crowder, R. G., & Melton, A. W. (1965). The Ranschburg phenomenon: Failures of immediate recall correlated with repetition of elements within a stimulus. *Psychonomic Science*, *2*, 295–296. Retrieved from papers2://publication/uuid/9C822265-52D4-4F11-AB61-F8E9BAA0545A

Cuddy, L. J., & Jacoby, L. L. (1982). When Forgetting Helps Memory : An Analysis of Repetition Effects. *Journal of Verbal Learning and Verbal Behavior*, *21*(4), 451–

467.

Davis, H., Rosner, T. M., D'Angelo, M. C., MacLellan, E., & Milliken, B. (2018). The role of conflict monitoring in the effect of selective attention on memory encoding. *Manuscript Submitted for Publication*.

Diemand-Yauman, C., Oppenheimer, D. M., & Vaughan, E. B. (2011). Fortune favors the bold (and the Italicized): Effects of disfluency on educational outcomes. *Cognition*, *118*(1), 111–5. http://doi.org/10.1016/j.cognition.2010.09.012

Ebbinghaus, H. (1913). *Memory: A Contribution to Experimental Psychology*. New York, NY: Teachers College, Columbia University (Original work published 1885).

Espinosa-García, M., Vaquero, J. M. M., Milliken, B., & Tudela, P. (2017). Recollection and familiarity for words and faces: a study comparing Remember–Know judgements and the Process Dissociation Procedure. *Memory*, *25*(1), 19–34. http://doi.org/10.1080/09658211.2015.1120310

Gellatly, A., Banton, P., & Woods, C. (1995). Salience and awareness in the Jacoby-Whitehouse effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(5), 1374–1379. http://doi.org/10.1037/0278-7393.21.5.1374

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, *91*(1), 1–67. http://doi.org/10.1037/0033-295X.91.1.1

Glanzer, M., & Adams, J. K. (1985). The mirror effect in recognition memory. *Memory & Cognition*, *13*(1), 8–20.

Gotts, S. J. (2016). Incremental learning of perceptual and conceptual representations and the puzzle of neural repetition suppression. *Psychonomic Bulletin and Review*, *23*(4), 1055–1071. http://doi.org/10.3758/s13423-015-0855-y

Greene, R. L. (1989). Spacing effects in memory: Evidence for a two-process account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(3), 371–377. http://doi.org/10.1037/0278-7393.15.3.371

Greene, R. L. (1991). The Ranschburg effect: The role of guessing strategies. *Memory & Cognition*, *19*(3), 313–7. http://doi.org/10.3758/BF03211155

Gregg, V. (1976). Word Frequency, Recognition and Recall. In J. Brown (Ed.), *Recall and Recognition*. Oxford, England: John Wiley & Sons.

Henson, R. N., & Gagnepain, P. (2010). Predictive, interactive multiple memory systems. *Hippocampus*, *20*(11), 1315–1326. http://doi.org/10.1002/hipo.20857

Henson, R. N., & Rugg, M. D. (2003). Neural response suppression, haemodynamic repetition effects, and behavioural priming. *Neuropsychologia*, *41*(3), 263–270. http://doi.org/10.1016/S0028-3932(02)00159-8

Higham, P. A., Perfect, T. J., & Bruno, D. (2009). Investigating strength and frequency effects in recognition memory using type-2 signal detection theory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *35*(1), 57–80. http://doi.org/10.1037/a0013865

Hintzman, D. L. (1974). Theoretical Implications of the Spacing Effect. In R. L. Solso (Ed.), *Theories in Cognitive Psychology: The Loyola Symposium* (pp. 77–99). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, *95*(4), 528–551.

http://doi.org/10.1037/0033-295X.95.4.528

Hirshman, E., & Mulligan, N. (1991). Perceptual interference improves explicit memory but does not enhance data-driven processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(3), 507–513.

Hirshman, E., Trembath, D., & Mulligan, N. (1994). Theoretical implications of the mnemonic benefits of perceptual interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(3), 608–620.

Hommel, B. (1998). Event files: Evidence for automatic integration of stimulus-response episodes. *Visual Cognition*, *5*(1–2), 183–216.

Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences*, *8*(11), 494–500. http://doi.org/10.1016/j.tics.2004.08.007

Jacoby, L. L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. *Journal of Verbal Learning and Verbal Behavior*, *17*(6), 649–667. http://doi.org/10.1016/S0022-5371(78)90393-6

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, *110*(3), 306–340. http://doi.org/10.1037/0096-3445.110.3.306

Jacoby, L. L., Kelley, C. M., & Dywan, J. (1989). Memory Attributions. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 391–422). Lawrence Erlbaum Associates, Inc.

Jacoby, L. L., & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology*, *118*(2), 126–135. http://doi.org/10.1037/0096-3445.118.2.126

Jacoby, L. L., Woloshyn, V., & Kelley, C. (1989). Becoming Famous Without Being Recognized: Unconscious Influences of Memory Produced by Dividing Attention. *Journal of Experimental Psychology: General*, *118*(2), 115–125.

Jahnke, J. C. (1969a). Output interference and the Ranschburg effect. *Journal of Verbal Learning and Verbal Behavior*, *8*, 614–621. http://doi.org/10.1016/S0022-5371(69)80113-1

Jahnke, J. C. (1969b). The Ranschburg Effect. *Psychological Review*, *76*(6), 592–605. http://doi.org/10.1037/h0028148

Jamieson, R. K., Holmes, S., & Mewhort, D. J. K. (2010). Global similarity predicts dissociation of classification and recognition: evidence questioning the implicit-explicit learning distinction in amnesia. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(6), 1529–1535. http://doi.org/10.1037/a0020598

Joordens, S., & Hockley, W. E. (2000). Recollection and familiarity through the looking glass: When old does not mirror new. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(6), 1534–1555.

Joordens, S., & Merikle, P. M. (1992). False recognition and perception without awareness. *Memory & Cognition*, *20*(2), 151–159. http://doi.org/10.3758/BF03197164

Kahana, M. J., & Howard, M. W. (2005). Spacing and lag effects in free recall of pure lists. *Psychonomic Bulletin & Review*, *12*(1), 159–164. http://doi.org/10.3758/BF03196362

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: object-specific integration of information. *Cognitive Psychology*, *24*(2), 175–219. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1582172

Kanwisher, N. G. (1987). Repetition blindness: Type recognition without token individuation. *Cognition*, *27*, 117–143.

Kelley, C. M., & Lindsay, D. S. (1993). Remembering Mistaken for Knowing: Ease of Retrieval as a Basis for Confidence in Answers to General Knowledge Questions. *Journal of Memory and Language*. http://doi.org/10.1006/jmla.1993.1001

Kelley, R., & Wixted, J. T. (2001). On the nature of associative information in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(3), 701–722. http://doi.org/10.1037/0278-7393.27.3.701

Kinoshita, S. (1995). The Word-Frequency Effect In Recognition Memory Versus Repetition Priming. *Memory & Cognition*, *23*(5), 569–580. http://doi.org/10.3758/bf03197259

Kinsbourne, M., & George, J. (1974). The mechanism of the word-frequency effect on recognition memory. *Journal of Verbal Learning and Verbal Behavior*, *13*(1), 63–69. http://doi.org/10.1016/S0022-5371(74)80031-9

Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences*, *4*(4), 138–147. http://doi.org/10.1016/S1364-6613(00)01452-2

Krebs, R. M., Boehler, C. N., De Belder, M., & Egner, T. (2015). Neural conflict-control mechanisms improve memory for target stimuli. *Cerebral Cortex*, *25*(3), 833–843. http://doi.org/10.1093/cercor/bht283

Kučera, H., & Francis, W. N. (1967). *Computational Analysis of Present-Day American English*. Providence, RI: Brown University Press.

Larsson, J., & Smith, A. T. (2012). fMRI repetition suppression: Neuronal adaptation or stimulus expectation? *Cerebral Cortex*, *22*(3), 567–576. http://doi.org/10.1093/cercor/bhr119

Law, M. B., Pratt, J., & Abrams, R. A. (1995). Color-based inhibition of return. *Perception & Psychophysics*, *57*(3), 402–408.

Lee, C. L. (1976). Short-term recall of repeated items and detection of repetitions in letter sequences. *Journal of Experimental Psychology: Human Learning & Memory*, *2*(2), 120–127. http://doi.org/10.1037/0278-7393.2.2.120

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, *87*(3), 252–271. http://doi.org/10.1037/0033-295X.87.3.252

McCabe, D. P., & Geraci, L. D. (2009). The influence of instructions and terminology on the accuracy of remember-know judgments. *Consciousness and Cognition*, *18*(2), 401–413. http://doi.org/10.1016/j.concog.2009.02.010

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. http://doi.org/10.3758/PBR.15.2.237

McGlone, M. S., & Tofighbakhsh, J. (2000). Birds of a Feather Flock Conjointly (?): Rhyme as Reason in Aphorisms. *Psychological Science*, *11*(5), 424–428.

Merikle, P. M., & Joordens, S. (1997). Parallels between Perception without Attention and Perception without Awareness. *Consciousness and Cognition*, *6*(6), 219–236. http://doi.org/10.1006/ccog.1997.0310

Mickes, L., Wixted, J. T., & Wais, P. E. (2007). A direct test of the unequal-variance signal detection model of recognition memory. *Psychonomic Bulletin & Review*, *14*(5), 858–865. http://doi.org/10.3758/BF03194112

Milliken, B., Joordens, S., Merikle, P. M., & Seiffert, A. E. (1998). Selective attention: a reevaluation of the implications of negative priming. *Psychological Review*, *105*(2), 203–229. http://doi.org/10.1037/0033-295X.105.2.203

Morey, R. D. (2008). Confidence intervals from normalized Data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*(2), 61–64.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, *16*(5), 519–533. http://doi.org/10.1016/S0022-5371(77)80016-9

Mulligan, N. W. (1996). The effects of perceptual interference at encoding on implicit memory, explicit memory, and memory for source. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(5), 1067–1087.

Mulligan, N. W. (1999). The effects of perceptual interference at encoding on organization and order: Investigating the roles of item-specific and relational information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(1), 54–69.

Mulligan, N. W., & Hirshman, E. (1995). Speed-accuracy trade-offs and the dual process model of recognition memory. *Journal of Memory and Language*, *34*, 1–18.

Nairne, J. S. (1988). The Mnemonic Value of Perceptual Identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(2), 248–255.

Nairne, J. S., Riegler, G. L., & Serra, M. (1991). Dissociative effects of generation on item and order retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(4), 702–709. http://doi.org/10.1037/0278-7393.17.4.702

Neill, W. T., & Mathis, K. M. (1998). Transfer-Inappropriate Processing: Negative Priming and Related Phenomena. In D. L. Mendin (Ed.), *The Psychology of Learning and Motivation (Vol. 38)* (pp. 1–44). San Diego, CA: Academic Press.

Olds, J. M., & Westerman, D. L. (2012). Can fluency be interpreted as novelty? Retraining the interpretation of fluency in recognition memory. *Journal of Experimental Psychology: Learning Memory, and Cognition*, *38*(3), 653–664. http://doi.org/10.1037/a0026784

Oppenheimer, D. M., & Frank, M. C. (2008). A rose in any other font would not smell as sweet: Effects of perceptual fluency on categorization. *Cognition*, *106*(3), 1178–1194. http://doi.org/10.1016/j.cognition.2007.05.010

Parks, C. M. (2013). Transfer-appropriate processing in recognition memory: Perceptual and conceptual effects on recognition memory depend on task demands. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(4), 1280–1286. http://doi.org/10.1037/a0030911

Parks, C. M., & Yonelinas, A. P. (2007). Moving beyond pure signal-detection models: comment on Wixted (2007). *Psychological Review*, *114*(1), 188-202; discussion

203-209. http://doi.org/10.1037/0033-295X.114.1.188

Peirce, J. W. (2007). PsychoPy-Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13. http://doi.org/10.1016/j.jneumeth.2006.11.017

Peirce, J. W. (2009). Generating Stimuli for Neuroscience Using PsychoPy. *Frontiers in Neuroinformatics*, *2*(January), 10. http://doi.org/10.3389/neuro.11.010.2008

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In *Attention and Performance X: Control of Language Processes* (pp. 531–556). http://doi.org/10.1162/jocn.1991.3.4.335

Quamme, J. R., Yonelinas, A. P., & Norman, K. A. (2007). Effect of unitization on associative recognition in amnesia. *Hippocampus*, *17*(9), 192–200. http://doi.org/10.1002/hipo

Rajaram, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory & Cognition*, *21*(1), 89–102.

Ratcliff, R. (1990). Connectionist models of recognition memory: Constraints imposed by learning and forgetting functions. *Psychological Review*, *97*(2), 285–308. http://doi.org/10.1037/0033-295X.97.2.285

Ratcliff, R., Sheu, C.-F., & Gronlund, S. D. (1992). Testing global memory models using ROC. *Psychological Review*, *99*(3), 518–535.

Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of Perceptual Fluency on Affective Judgments. *Psychological Science*, *8*(1), 45–48.

Roediger, H. L., & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, *1*(3), 181–210. http://doi.org/10.1111/j.1745-6916.2006.00012.x

Rosner, T. M., D’Angelo, M. C., MacLellan, E., & Milliken, B. (2015). Selective attention and recognition: effects of congruency on episodic learning. *Psychological Research*, *79*(3), 411–424. http://doi.org/10.1007/s00426-014-0572-6

Rosner, T. M., Davis, H., & Milliken, B. (2015). Perceptual blurring and recognition memory: A desirable difficulty effect revealed. *Acta Psychologica*, *160*, 11–22. http://doi.org/10.1016/j.actpsy.2015.06.006

Rosner, T. M., Lopez-Benitez, R., D’Angelo, M. C., Thomson, D. R., & Milliken, B. (2018). Remembering “primed” words: A counter-intuitive effect of repetition on recognition memory. *Canadian Journal of Experimental Psychology*, *72*(1), 24–37. http://doi.org/10.1037/CEP0000139

Rotello, C. M., Macmillan, N. A., Reeder, J. A., & Wong, M. (2005). The remember response: subject to bias, graded, and not a process-pure indicator of recollection. *Psychonomic Bulletin & Review*, *12*(5), 865–73. http://doi.org/10.3758/BF03196778

Rotello, C. M., Macmillan, N. A., & Van Tassel, G. (2000). Recall-to-Reject in Recognition: Evidence from ROC Curves. *Journal of Memory and Language*, *43*(1), 67–88. http://doi.org/10.1006/jmla.1999.2701

Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance*, *3*(1), 1–17. http://doi.org/10.1037/0096-1523.3.1.1

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon.

Journal of Experimental Psychology: Human Learning & Memory, *17*(6), 359–369. http://doi.org/10.3758/BF03198475

Slotnick, S. D., & Dodson, C. S. (2005). Support for a continuous (single-process) model of recognition memory and source memory. *Memory & Cognition*, *33*(1), 151–70. http://doi.org/10.3758/BF03195305

Spadaro, A., He, C., & Milliken, B. (2012). Response to an intervening event reverses nonspatial repetition effects in 2AFC tasks: nonspatial IOR? *Attention, Perception & Psychophysics*, *74*(2), 331–49. http://doi.org/10.3758/s13414-011-0248-x

Spadaro, A., & Milliken, B. (2013). Subjective expectancy and inhibition of return: A dissociation in a non-spatial two-alternative forced choice task. *Psycológica*, *34*, 199–219.

Stretch, V., & Wixted, J. T. (1998). On the difference between strength-based and frequency-based mirror effects in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(6), 1379–1396. http://doi.org/10.1037/0278-7393.24.6.1379

Summerfield, C., Trittschuh, E. H., Monti, J. M., Mesulam, M., & Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nature Neuroscience*, *11*(9), 1004–1006. http://doi.org/10.1038/nn.2163

Summerfield, C., Wyart, V., Johnen, V. M., & Gardelle, V. De. (2011). Human scalp electroencephalography reveals that repetition suppression varies with expectation. *Frontiers in Human Neuroscience*, *5*(July), 1–13. http://doi.org/10.3389/fnhum.2011.00067

Thorndike, E. L., & Lorge, I. (1944). *The teacher's handbook of 30,000 words.* New York: Columbia University.

Tipper, S. P. (1985). The negative priming effect: inhibitory priming by ignored objects. *The Quarterly Journal of Experimental Psychology*, *37*(4), 571–590. http://doi.org/10.1080/14640748508400920

Underwood, B. J. (1969). Some correlates of item repetition in free-recall learning. *Journal of Verbal Learning and Verbal Behavior*, *8*(1), 83–94. http://doi.org/10.1016/S0022-5371(69)80015-0

Van Selst, M., & Jolicoeur, P. (1994). A solution to the effect of sample size on outlier elimination. *The Quarterly Journal of Experimental Psychology*, *47*(3), 631–650. http://doi.org/10.1080/14640749408401131

Westerman, D. L. (2008). Relative fluency and illusions of recognition memory. *Psychonomic Bulletin & Review*, *15*(6), 1196–1200. http://doi.org/10.3758/PBR.15.6.1196

Whittlesea, B. W. A. (1993). Illusions of Familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(6), 1235–1253.

Whittlesea, B. W. A. (1997). Production, Evaluation, and Preservation of Experiences: Constructive Processing in Remembering and Performance Tasks. In *Psychology of Learning and Motivation - Advances in Research and Theory* (Vol. 37, pp. 211–264). Elsevier Masson SAS. http://doi.org/10.1016/S0079-7421(08)60503-4

Whittlesea, B. W. A., Jacoby, L. L., & Girard, K. (1990). Illusions of immediate memory: Evidence of an attributional basis for feelings of familiarity and perceptual quality.

Journal of Memory and Language, *29*(6), 716–732. http://doi.org/10.1016/0749-596X(90)90045-2

Whittlesea, B. W. A., & Williams, L. D. (1998). Why do strangers feel familiar, but friends don't? A discrepancy-attribution account of feelings of familiarity. *Acta Psychologica*, *98*, 141–165. http://doi.org/10.1016/S0001-6918(97)00040-1

Whittlesea, B. W. A., & Williams, L. D. (2000). The Source of Feelings of Familiarity: The Discrepancy-Attribution Hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(3), 547–565. http://doi.org/10.1037//0278-7393.26.3.547

Whittlesea, B. W. A., & Williams, L. D. (2001). The Discrepancy-Attribution Hypothesis: I. The Heuristic Basis of Feelings of Familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(1), 3–13.

Witherspoon, D., & Allan, L. G. (1985). The effect of a prior presentation on temporal judgments in a perceptual identification task. *Memory & Cognition*, *13*(2), 101–111. http://doi.org/10.3758/BF03197003

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review*, *114*(1), 152–176. http://doi.org/10.1037/0033-295X.114.1.152

Wixted, J. T., & Mickes, L. (2010). A continuous dual-process model of remember/know judgments. *Psychological Review*, *117*(4), 1025–1054. http://doi.org/10.1037/a0020874

Wixted, J. T., & Stretch, V. (2004). In defense of the signal detection interpretation of remember/know judgments. *Psychonomic Bulletin & Review*, *11*(4), 616–641. http://doi.org/10.3758/BF03196616

Wood, T. J., & Milliken, B. (1998). Negative priming without ignoring. *Psychonomic Bulletin & Review*, *5*(3), 470–475. http://doi.org/10.3758/BF03208824

Xue, G., Mei, L., Chen, C., Lu, Z.-L., Poldrack, R., & Dong, Q. (2011). Spaced Learning Enhances Subsequent Recognition Memory by Reducing Neural Repetition Suppression. *Journal of Cognitive Neuroscience*, *23*(7), 1624–1633. http://doi.org/10.1162/jocn.2010.21532

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(5), 601–621.

Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory - evidence for a duel-process model. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *20*(6), 1341–1354. http://doi.org/10.1037/0278-7393.20.6.1341

Yonelinas, A. P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. *Memory & Cognition*, *25*(6), 747–763. http://doi.org/10.3758/BF03211318

Yonelinas, A. P. (2001). Consciousness, control, and confidence: the 3 Cs of recognition memory. *Journal of Experimental Psychology: General*, *130*(3), 361–379. http://doi.org/10.1037/0096-3445.130.3.361

Yonelinas, A. P. (2002). The Nature of Recollection and Familiarity: A Review of 30

Years of Research. *Journal of Memory and Language*, *46*(3), 441–517. http://doi.org/10.1006/jmla.2002.2864

Yonelinas, A. P., Aly, M., Wang, W. C., & Koen, J. D. (2010). Recollection and familiarity: Examining controversial assumptions and new directions. *Hippocampus*, *20*(11), 1178–1194. http://doi.org/10.1002/hipo.20864

Yonelinas, A. P., Dobbins, I., Syzmanski, M. D., Dhaliwal, H. S., & King, L. (1996). Signal-detection, threshold, and dual-process models of recognition memory: ROCs and conscious recollection. *Consciousness and Cognition*, *5*(4), 418–441. http://doi.org/10.1006/ccog.1996.0026

Yonelinas, A. P., & Jacoby, L. L. (1995). The Relation between Remembering and Knowing as Bases for Recognition: Effects of Size Congruency. *Journal of Memory and Language*, *34*(5), 622–643.

Yonelinas, A. P., & Parks, C. M. (2007). Receiver operating characteristics (ROCs) in recognition memory: a review. *Psychological Bulletin*, *133*(5), 800–32. http://doi.org/10.1037/0033-2909.133.5.800

Yue, C. L., Castel, A. D., & Bjork, R. A. (2013). When disfluency is--and is not--a desirable difficulty: The influence of typeface clarity on metacognitive judgments and memory. *Memory & Cognition*, *41*(2), 229–241. http://doi.org/10.3758/s13421-012-0255-8

Zhao, X., Wang, C., Liu, Q., Xiao, X., Jiang, T., Chen, C., & Xue, G. (2015). Neural mechanisms of the spacing effect in episodic memory: A parallel EEG and fMRI study. *Cortex*, *69*, 76–92. http://doi.org/10.1016/j.cortex.2015.04.002