2017

Would disfluency by any other name still be disfluent? Examining the boundary conditions of the disfluency effect

Jason Geller
Iowa State University

Would disfluency by any other name still be disfluent?
Examining the boundary conditions of the disfluency effect

by

Jason Geller

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Psychology

Program of Study Committee
Veronica J Dark, Major Professor
Mary L. Still
Shana Carpenter
Kristi Costabile
Marcus Crede

Iowa State University

Ames, Iowa

2017

TABLE OF CONTENTS

ACKNOWLEGEMENTS

First and foremost, I would like to extend my deepest gratitude to my two major professors: Veronica J. Dark and Mary L. Still. When Alison passed away two years ago, I honestly did not know if I was going to be able to continue on in the program. Both of you took on the responsibility and burden of mentoring a graduate student (me), when you had no obligation to do so. Working with both of you has been a highlight of my graduate school experience.

Veronica, although you and I had, at times, an acrimonious relationship, I have gleaned a lot from you about research. You have also contributed greatly to my professional development. The hours we spent talking about my data were invaluable. I hope I can put on my Veronica hat throughout my career.

Mary, although you probably would not agree, you are very much like Alison in your mentoring style. You are caring, but also critical of my work. You expect nothing but the best from me, but were never without encouragement. You have made me a better writer and researcher. Having a colleague to work with on word nerd projects with has kept me sane over the last few years.

I would also like to thank Shana. S. Carpenter. I want to thank you for the all the opportunities and support you have given me over the past year. I have learned a great deal from you. Our long conversations about research were invaluable and I will miss them.

Last, but not least, I would like to thank my best friend and significant other, Kelly Ann Kane. You have been an incredible source of inspiration to me throughout this process. Your ardor for writing has pushed me to be a better writer, and I hope that is reflected herein. I have greatly benefited from us sharing ideas with one another and our brainstorming sessions. Further,

whenever I needed help with writing, or a reprieve, you were always there. If it was not for our long hours spent together in random places around Ames and upstate NY (e.g., Dark Lab, Wegmans, Ames Library, Keuka Lake), I don't think I could have finished this monolithic undertaking.

ABSTRACT

When exposed to words presented under perceptually disfluent conditions (e.g., words written in **Haettenschweil** font), participants have difficulty recognizing the words. Those same words, though, may be better remembered than words presented in standard type font. This counterintuitive finding is referred to as the disfluency effect. Evidence for this disfluency effect, however, has been mixed. Using a recognition memory task, I examined five variables that may contribute to the inconsistent findings: type of judgments of learning (JOLs), encoding instructions, type of disfluency manipulation, encoding duration, and retention interval between study and test. Experiment 1 employed a masking manipulation and examined the influence of type of JOLs (item-by-item JOLs vs. aggregate JOLs) along with encoding instructions (incidental vs. intentional). Experiments 2 and 3 explored the locus of the disfluency effect by examining perceptual disfluency manipulations that tap different loci during word recognition: low-level blurring (pre-lexical) and cursive handwriting (lexical). Experiment 4 examined the role of encoding duration on the disfluency effect. Experiment 5 examined whether list design (blocked vs. mixed) moderates the disfluency effect. Experiment 6 examined whether the benefits of disfluency extend over longer durations (24 hours). Results across the six experiments indicated that the disfluency effect is modulated by testing expectancy, type of disfluency manipulation, and encoding duration. A disfluency effect was observed only under incidental instructions with a sufficiently long encoding duration. Further, I found that a pre-lexical manipulation (i.e., blurring) did not produce a disfluency effect, but a lexical perceptual disfluency manipulation (i.e., cursive) did. This cursive disfluency effect was moderated by legibility: easy-to-read cursive words tended to be better remembered than hard-to-read cursive

words. This finding was bolstered by a meta-analysis. Taken together, these results challenge extant accounts of the disfluency effect. The research comprising this dissertation furthers the theoretical understanding of the disfluency effect as well at its practical implications.

CHAPTER 1. BACKGROUND LITERATURE

Introduction

In the educational milieu, students are confronted with a multitude of decisions while studying. Students must decide what to study, when to study, how long to study, and what techniques to use. One hallmark of a self-regulated learner, then, is the ability to accurately monitor and regulate information that is processed—that is, to determine what one does and does not know and allocate attentional resources appropriately (Bjork, Dunlosky, & Kornell, 2013). Being able to accurately monitor learning is an essential metacognitive skill that promotes positive learning outcomes (e.g., increased comprehension, stronger memory representations, and better grades). For instance, while studying for an exam, students who accurately monitor their learning could identify when they do not have mastery over the material and decide to study more thoroughly the information they have not yet mastered. Conversely, students who inaccurately monitor their learning might decide to terminate studying prematurely, having not fully mastered the material, which would hinder learning. One pervasive cue that influences these metacognitive judgments is fluency: the subjective ease associated with processing information (Alter & Oppenheimer, 2009; Koriat & Bjork, 2006).

In the domain of learning and memory, an important question to ask is whether fluency is an adaptive cue—that is, does a person's subjective (predicted) confidence in his or her memory correlate positively with objective (actual) memory? Some studies have shown a positive correlation between subjective confidence and actual performance (e.g., Begg, Duft, Lalonde, Melnick, & Sanvito, 1988; Koriat, Ma'ayan, & Nusinson, 2006). However, it is most often the case that fluency elicits overconfidence (predicted memory > actual memory) in assessments of learning, which makes fluency a maladaptive learning cue (Fiedler, 2013).

The discrepancy in the effect of fluency on subjective and objective assessments of memory was termed a *metacognitive illusion* by Rhodes and Castel (2008). This metacognitive illusion has been shown for a variety of different fluency manipulations. For example, items that are bigger (vs. smaller), items spoken loudly (vs. softly), or items that are studied while holding something heavy (vs. light) are judged as more memorable, but there is not any mnemonic benefit (Alban & Kelley, 2013; Rhodes & Castel, 2009). One way to mitigate this metacognitive illusion is to introduce difficulties into the learning process by making encoding less fluent. Manipulations that introduce obstacles into the learning process, but that also have the added effect of improving long-term retention, have been conceptualized as *desirable difficulties* (Bjork, 1994; McDaniel & Butler, 2010). Notable examples of desirable difficulties include having participants generate information from word fragments instead of passively reading intact words (e.g., Slamecka & Graf, 1978), spacing out study sessions in time and space instead of massing study sessions (e.g., Baddely & Longman, 1978), and having participants engage in retrieval practice after studying instead of simply restudying the information (Roediger & Karpicke, 2006).

Other manipulations that may impart desirable difficulty involve changing the perceptual characteristics of stimuli to make them harder to identify. For example, providing individuals with text that is in an atypical font that is difficult to read can lead to increased memory for that text compared to an easy to read text (e.g., Diemand-Yaumen, Oppenheimer, & Vaughan, 2011). The benefit associated with processing difficulty brought forth by changing perceptual characteristics of a text- the *disfluency effect* - is quite surprising as many traditional models of memory posit that encoding operates within a limited capacity channel (e.g., Atkinson & Shiffrin, 1968; Baddeley, 1981). In these models, increased demands on working memory would increase cognitive load, and therefore should be more likely to hurt rather than to help memory. If working memory capacity is limited, perceptual disfluency should be undesirable for learning.

Before proceeding with a discussion of disfluency, it is important to clarify two different uses of the term “disfluency.” It can be used to describe a stimulus manipulation or to describe the impact of that manipulation on memory. To aid clarity, the following terminology will be used throughout this dissertation. The term *perceptual disfluency* will be used to refer to any manipulation that makes processing more difficult whereas *disfluency effect* refers to any memory benefit evoked by perceptual disfluency.

Although the disfluency effect has been demonstrated in a variety of experiments (e.g., Diemand-Yaumen et al., 2010; French et al., 2013; Sungkhasette, Friedman, & Castle, 2011), there have also been several experiments that failed to find the effect (e.g., Magreehan, Serra, Schwartz & Narciss, 2015; Rhodes & Castel, 2008, 2009; Rummer, Scheweppe, & Schewede, 2016; Yue et al., 2013). Much of the literature examining the disfluency effect consists of conceptual replications using a wide variety of manipulations and tasks; this variety makes it difficult to know whether a replication fails due to procedural differences or due to the absence of a true disfluency effect. Given the equivocal findings in the literature, the current research aims to identify important moderating or boundary conditions of the effect (Kuhl & Eitel, 2016; Oppenheimer & Alter, 2014). The present series of experiments is designed to test the impact of type of JOLs, encoding instructions, type of perceptual manipulations, encoding duration, and retention interval on the disfluency effect. Manipulating variables systematically will not only elucidate the conditions under which perceptual disfluency does and does not lead to better memory, but should also shed light on past failures to replicate the disfluency effect.

The Disfluency Effect

Traditionally, the disfluency effect has been examined using a simple study-test paradigm. During study, stimuli are presented either perceptually unaltered (fluent) or perceptually manipulated (disfluent). In the earliest research, disfluency was produced by presenting stimuli briefly (approximately 100 ms) and backward masking them, typically with a row of hash marks, while fluent stimuli were presented unmasked and for a longer duration (e.g., Hirshman & Mulligan, 1991; Hirshman et al., 1994; Mulligan, 1996; Nairne, 1988). The backward mask is assumed to interrupt perceptual processing of the stimulus, thereby making identification more difficult, but not impossible (accuracy for both conditions is high). To ensure that both the masked and unaltered stimuli were identified accurately, participants named stimuli immediately after presentation. At test, participants' memory is better for disfluent items compared to the fluent items. Better memory for such perceptually disfluent stimuli has been obtained in a number of different explicit memory tests: free recall (e.g., Hirshman & Mulligan, 1991), recognition (e.g., Narnie, 1988), and cued-recall (e.g., Mulligan, 1996).

More recent research has examined the disfluency effect using a variety of other perceptual disfluency manipulations. For example, Sungkhasette et al. (2011) found that presenting stimuli in an inverted orientation compared to an upright orientation during encoding produced better retention of the material in a free recall test. Moreover, in a highly publicized paper with more than 150 citations (Kuhl & Eitel, 2016), Diemand-Yauman et al. (2010) examined the effects of disfluency in both the laboratory and classroom. In Experiment 1, participants studied a list of characteristics associated with space aliens in either a disfluent typescript (*Comic Sans MS* and ***Bodoni MT***) or a more common, fluent typescript (*Arial*). At test, participants showed better memory on a cued-recall task for characteristics presented in a disfluent typescript compared to a fluent typescript (see French et al., 2013, and Weltman & Eakin, 2014, for replications of this effect). In Experiment 2, Diemand-Yauman et al. manipulated the typescript (*Monotype Corsiva*, *Comic Sans Italicized*, and **Haettenschaller**) of PowerPoint slides and handouts for a single unit in advanced placement (AP) English, History, Chemistry, or Physics. Similar to Experiment 1, the disfluent typescripts resulted in better memory than the fluent typescripts on a unit exam.

Failures to Replicate the Disfluency Effect

The preceding discussion suggests that the disfluency effect is a true effect in that better memory was obtained due to the perceptual disfluency. However, there also have been some failures to replicate the disfluency effect. The fact that there have been several failures to replicate the disfluency effect raises questions about whether perceptual disfluency is generally useful as a desirable difficulty. In a simple memory paradigm, for instance, Rhodes and Castel (2008) showed that words in a smaller sized font (18 point) were judged as being more disfluent compared to words printed in a larger sized font (48 point), but the larger font did not lead to better memory —recall differences between the smaller and larger size fonts were negligible (see Kornell et al., 2011; Mueller, Dunlosky, Tauber, & Rhodes, 2014; Susser, Mulligan, & Besken, 2013, for similar failures to replicate). In a more recent study, Yue et al. (2013) examined the disfluency effect for blurred words. They examined the effect of blurring across several factors: type of task (recall vs. recognition), study duration (500 ms vs. 2 s), and design (within- vs. between-item lists). Just as with Rhodes and Castle (2008), there was no memory benefit for disfluent stimuli. Failures to replicate the disfluency effect have also been found with atypical or hard-to-read fonts (e.g., Magreehan et al., 2015) and hard to hear auditory information (e.g., Rhodes & Castel, 2009).

Failures to replicate are not only limited to simple memory paradigms. Eitel, Kuhl, Scheiter, and Gerjets (2014), for example, showed across multiple experiments that having participants study multimedia discourse (i.e., pictures embedded within text) on the mechanisms of a toilet flush in a disfluent font and/or low-quality photocopy did not engender greater memory for the material on a comprehension test. Similarly, Strukelj, Scheiter, Nystrom, and Holmqvist (2015) found that reading a disfluent discourse (i.e., words that were passed through a low-level filter) on how airplanes achieve lift did not result in better memory for material in a free recall task. Lastly, Rummer, Schweppe, and Schewede (2016) attempted a conceptual replication of Diemand-Yauman et al. (2010) by examining distinctiveness as a possible mechanism of the disfluency effect. Rummer et al. did not find a disfluency effect, thus, no evidence supporting distinctiveness as a memory-enhancing mechanism.

The fact that perceptual disfluency does not always promote positive learning outcomes raises serious concerns regarding disfluency as a desirable difficulty. When a finding fails to replicate, it is important to examine why studies failed to find an effect, if it is true. To better aid in understanding when perceptual disfluency is desirable and when it is not, theoretical accounts of the disfluency effect are discussed next.

Theoretical Accounts of the Disfluency Effect

Making a word perceptually disfluent can sometimes enhance memory for that word. This finding leads to the question: When and how does the disfluency effect arise during the course of word processing? Each of the accounts discussed in the following paragraphs propose that disfluency arises during the processing of the word. Where they differ is in their explanation of how and when the disfluency effect arises during processing.

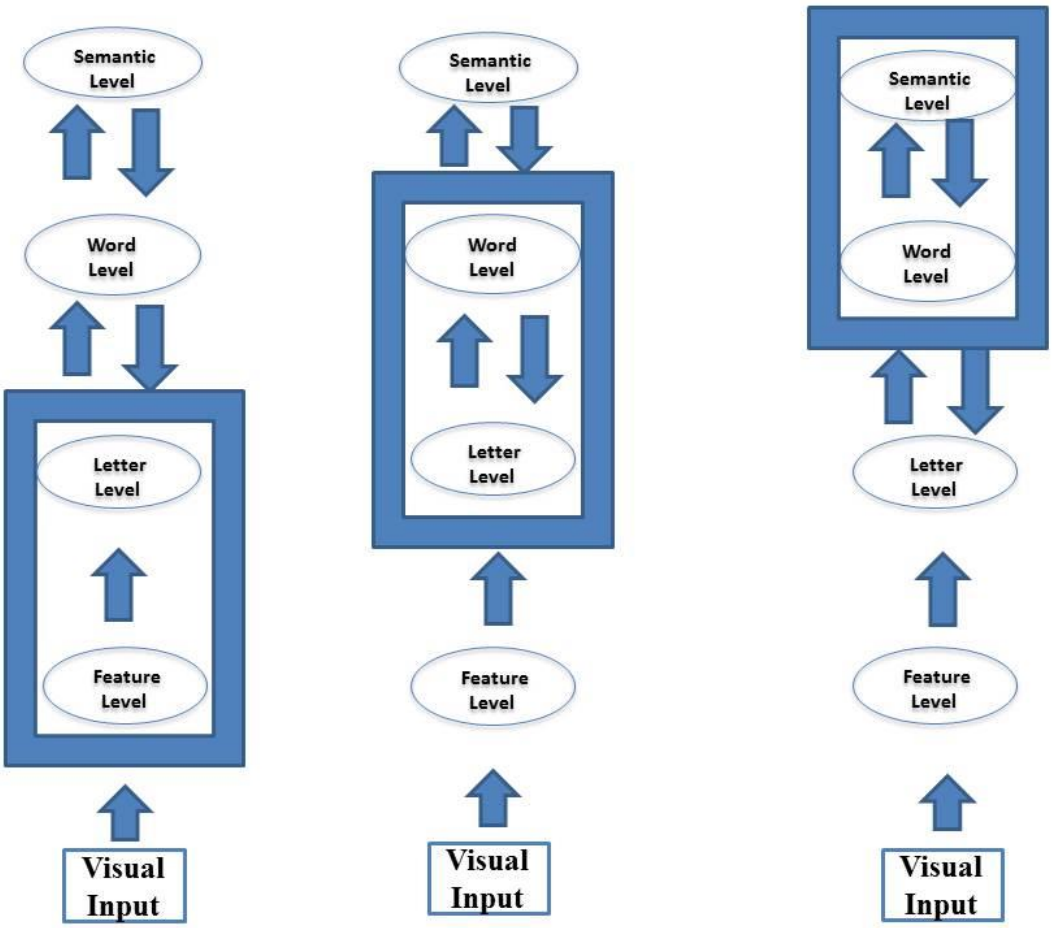

Figure 1. The four levels of representation used during the course of word recognition. The box on the left represents a pre-lexical locus, with perceptual disfluency increasing the amount of noise in the activation from the feature- level to letter-level. The box in the middle depicts a lexical locus, with perceptual disfluency not only increasing the amount of noise in the activation from the feature-level to letter-level, but also increasing the amount of feedback from the word-level to the letter-level. The box on the right depicts a post-lexical locus, with perceptual disfluency increasing the amount of feedback from the semantic-level to word-level as a result of disfluency experienced during word processing.

To better explore the existing explanations of the disfluency effect for words as well as the loci of the effect, I have created a descriptive model (see Figure 1) that depicts four levels of representation (feature, letter, word, and semantic) involved in word recognition. Each level is represented by an ellipse. Within the model, the effect of perceptual disfluency on memory can arise at one of three loci: pre-lexical, lexical, and post-lexical. In other words, the effect of disfluency on memory can occur before word recognition occurs, during word recognition, or after the word has already been recognized. The loci are represented by rectangles. This model is based on the widely-used computational model of visual word recognition, the Interactive-Activation (IA) model (McClelland & Rumelhart, 1981).

Within the IA model (McClelland & Rumelhart, 1981), word recognition is the result of complex facilitatory and inhibitory interactions between representations at the first three levels: features, letters, and words (Davis, 2003; Lupker & Davis, 2009; Morris & Still, 2010). In Figure 1, I have appended a semantic-level of representation to accommodate theories of the disfluency effect in which context or known associations (i.e., semantic information) play a part. Within the IA model, each of the levels becomes activated in parallel. When an individual reads a word, processing begins in a bottom-up fashion with facilitatory activation feeding forward from features to letters to words. As activation spreads within each representational level, it is inhibitory; inhibition is most pronounced at the word-level, where representations for the presented word must compete with orthographically similar words (Coltheart., Davelaar, Jonasson, & Besner, 1977). For example, if *CAT* were the presented word, its representation would compete with the representations of *BAT*, *FAT*, *CAR*, *CAP*, *COT*, *CUT*, and others. The representation with the most activation "beats down" its competitors and reaches a critical threshold, resulting in, what Balota (1990) dubbed the "magic moment", or word recognition.

An important feature of activation in IA-like models is that it is not only bottom up, but is also top down. That is, the letter-level not only feeds forward to the word level, but the word level also feeds back down to the letter level to aid in word recognition. The interactive nature of activation between letters and words is illustrated by the word superiority effect wherein individuals are better able to identify letters when they are presented in words than when those same letters are presented in isolation (Cattell, 1886; Richer, 1969). The word superiority effect occurs because letters within words get top down, word-level to letter-level feedback, resulting in better identification for letters presented in words. Letters presented in isolation only receive bottom-up, letter-level activation, so the identification task does not benefit from word-level feedback as letters do not have word-level representations.

Within the model presented in Figure 1, the assumption is that perceptual disfluency would produce a difficulty in processing by introducing considerable perceptual "noise", or activation variability, during the word recognition process. It is the additional processing needed to overcome the noise that creates a memory advantage for disfluent stimuli. The additional processing producing the memory advantage could arise at any of these stages of processing. In all cases, the perceptual disfluency introduces noise at the pre-lexical level. The disfluency effect could arise from this additional processing at the pre-lexical locus. Or, the disfluency effect could arise when additional processing also is required at the lexical locus, with increased interactivity from word to letter levels being responsible for the memory benefit. Or, the disfluency effect could arise when there is additional processing involving the post-lexical locus. These three loci are discussed next.

Pre-lexical locus. All word processing proceeds through the loci depicted in Figure 1. However, when words are perceptually disfluent, processing at some loci may be influenced to a greater degree than others. A perceptual disfluency manipulation that exerts an effect at a pre-lexical locus causes activation variability or noise from the feature level to the letter level. This variability in mapping features to letters slows down letter recognition, but leaves the other loci (lexical and post-lexical) unaffected. It is the requirement for additional processing in order to determine what the letters are that results in the memory benefit for disfluent items. Hirshman et al. (1994, Experiment 5) examined this pre-lexical locus account by varying the perceptual difficulty of features via a luminance manipulation in which words were presented in dark-grey vs. white on a black background. Reducing the contrast should make letters harder to read. Hirshman et al. reasoned that, if the pre-lexical account is correct, there should be a disfluency effect with better memory for words presented in the disfluent condition (dark grey) than for words presented in the fluent condition (white); however, both fluent and disfluent conditions produced comparable recall rates, thereby providing evidence against the pre-lexical locus account of the disfluency effect.

Lexical locus. A perceptual disfluency manipulation that exerts an effect on a lexical locus not only increases activation variability or noise from the feature to the letter level, but also increases the amount of feed forward activation from feature to letter levels and the amount of feed backward activation from word to letter levels. One lexical account, the compensatory-processing account (Hirshman et al., 1994; Mulligan, 1996), ties the disfluency effect directly to this increased lexical processing brought forth by the perceptual disfluency manipulation. Manipulations affecting a lexical locus produce a qualitative change in word processing that requires more top-down, word-level to letter-level, processing. That is, perceptual disfluency manipulations affecting the lexical locus not only slow down processing, but change the type of processing a word undergoes. It is this increased lexical processing that enhances memory for disfluent stimuli.

One perceptual disfluency manipulation that affects this lexical locus is masking. Results from computational modeling have shown that masking a word facilitates increased lexical feedback from the word to letter level during processing (McCelland & Rumelhart, 1981). This is the basis for the word superiority effect discussed before (Cattell, 1886; Richer, 1969). According to this lexical, or compensatory processing (Mulligan, 1996), account of the disfluency effect, perceptual disfluency manipulations that do not affect the lexical locus strongly should not produce a disfluency effect. Indeed, Westerman and Greene (1997) showed that nonwords fail to elicit the disfluency effect. Nonwords do not have lexical representations, and thus cannot benefit from the lexical feedback that is the basis of the compensatory-processing account. Similarly, Mulligan (1998) and Westerman and Greene (1998) showed that masking a stimulus that activates the lexical level very weakly (e.g., low imageability or low frequency words) does not produce the disfluency effect.

Post-lexical locus. Lastly, a perceptual manipulation may affect a post-lexical locus. Extant post-lexical accounts of the disfluency effect (e.g., Alter, 2013; Alter, Oppenheimer, Epley, & Eyre, 2007; Diemand-Yaumen et al., 2010) suggest that the memory benefit from perceptual disfluency arises at a post-lexical locus due to increased feedback from the semantic level to the word level. The memorial benefit for perceptually disfluent words is not a direct result of processing the noisy input; rather current post-lexical accounts suggest that the experience of disfluency during the process of word recognition functions as a metacognitive (subjective) cue to engage in semantic elaboration of the word. From this perspective, deeper, more semantic, processing is the mechanism that produces enhanced memory.

This is related to the traditional levels-of-processing framework proposed by Craik and Lockhart (1972). This framework suggests that strength of a memory is determined by the level of processing a stimulus undergoes. Words that are processed to deeper (i.e., semantic) levels will be better remembered than words processed at shallower (i.e., featural or phonetic) levels. For example, a researcher might induce deeper (semantic) processing by asking a participant if a word that is presented (e.g., *beagle*) is a part of a category (e.g., “Is it a type of dog?”); whereas a researcher might induce shallow processing by asking a participant if a presented word rhythms with another word (e.g., "Does *beagle* rhythm with *negal*?"). Words processed more deeply have a stronger memory representation. It is this deeper processing that is presumed to be the sole mechanism driving the disfluency effect (Alter et al., 2007).

A more contemporary description of this post-lexical account is framed in terms of two metaphorical systems: System 1 and System 2 (Evans, 2006; Kahneman, 2011). In this view, humans are thought to possess two modular information-processing systems, called System 1 and System 2. System 1 processing is characterized as automatic, intuitive, effortless, and fast—tthis system supports shallower processing of information. In contrast, System 2 processing, which involves the allocation of finite processing resources, is controlled, effortful, deliberate, and slow—this system supports deeper information processing. It is important to note that System 1 and System 2 interact with one another. That is, the automaticity of System 1 computations can be modulated or overridden by System 2 computations. When information processing is disfluent, it is suggested that System 1 recruits System 2 resources (Alter et al., 2007). This increased System 2 processing is thought to lead to the memory benefit for disfluent items.

The idea that the disfluency effect arises at a post-lexical locus as a result of deeper, more semantic processing is widely put forth as an explanation for the disfluency effect (Atler, 2013; Alter et al., 2007; Diemand-Yauman et al., 2010). However, Hirshman et al. (1994, Experiment 1) provided evidence against this explanation when they manipulated word study time. Hirshman et al. reasoned that as study time increases, so does the amount of higher-level (post-lexical) elaboration a perceptually disfluent item receives. Thus, if the disfluency effect is post-lexical, it should be influenced by study time. In their study, increasing study time did not increase the magnitude of the disfluency effect. Mulligan (1996, Experiments 1 and 3) further examined the post-lexical account by examining perceptual disfluency in the category-exemplar production task, an implicit memory task purportedly driven by higher-level (conceptual/semantic) processing. However, no disfluency effect was found with this implicit task. This suggests that the locus of the effect is not post-lexical.

Another post-lexical mechanism that may be involved in the disfluency effect is distinctiveness (see Schmidt, 1991, for a review). The distinctiveness account suggests that it is not the experience of disfluency during word recognition that acts a cue to process the word more fully. Instead, it is the subjective impression of atypicality or unusualness of the perceptually disfluent word that produces the better memory. To test this, Hirshman et al. (1994, Experiment 4) presented half of the items in an atypical color (green-on-brown background) compared to stimuli in a more typical color (white-on-black background). Although the items presented in the atypical color were more distinctive, the color manipulation did not affect memory. Additionally, Sungkhasette et al. (2011) examined the role of distinctiveness by dividing inverted stimuli presented in the study phase into two halves. Sungkhasette et al. reasoned that if the disfluency effect found for inverted items was due to distinctiveness, then the memory advantage should dissipate across the two halves of the experiment. Comparing recall performance for the two halves showed a disfluency effect that did not differ between the two halves, leading to the conclusion that that distinctiveness does not play a role in the disfluency effect.

As illustrated in the preceding paragraphs, there is evidence for and against the disfluency effect arising at all three loci: pre-lexical, lexical, and post-lexical. Given the mixed evidence, it is important to examine this more systematically as the conditions producing and not producing the disfluency effect have important theoretical implications.

Table 1

Studies Finding a Disfluency Effect: a Function of Manipulation, JOL and Testing Expectancy

Disfluency Manipulation	JOL	Testing Expectancy
MASKING		
Narnie (1988)	None	Incidental
Westerman & Greene (1997)	None	Incidental
Besken & Mullian (1991, Experiment 1)	Aggregate	Intentional
Hirshman & Mulligan (1991, Experiment 1-2)	None	Incidental
Mulligan (1996, Expeirment 1)	None	Incidental
Mulligan (1999, Experiment 1-3)	None	Incidental
Hirshman, Trembath, & Mulligan (1994, Experiment 1-2)	None	Incidental
INVERSION		
Sungkhasettee, Freidman, & Castel (2011)	Aggregate	Intentional
OTHER		
Besken & Mulligan (2014) – Auditory missing letter		
Experiment 1	Aggregate	Intentional
Experiments 2-3	Item	Intentional
Susser, Mulligan, & Besken (2013) – Transposed letters	Aggregate	Intentional

Table 2

Studies Finding No Disfluency Effect: a Function of Manipulation, JOL and Testing Expectancy

Disfluency Manipulation	JOL	Testing Expectancy
MASKING		
Besken & Mullian (2013, Experiment 2)	Item	Intentional
Hirshman & Mulligan (1991)	None	Incidental
Hirshman, Trembath, & Mulligan (1994)	None	Incidental
COLOR CONTRAST		
Hirshman et al. (1994, Experiment 4)	None	Incidental
BLURRING/FILTERING		
Yue, Castel, & Bjork (2013, Experiment 1A, 2A, 3)	Item	Intentional
Hirshman et al. (1994, Experiment 5)	None	Incidental
FONT STYLE		
Magreehan, Serra, Schwartz, & Narciss (2015)	Item	Intentional
Jia, Li, Li, Zhang, Cao, & Li (2015)	Item	Intentional
Rhodes & Castel (2008, Experiment 6)	Item	Intentional
FONT SIZE		
Rhodes & Castel (2008, Experiments 1-5)	Item	Intentional
Mueller, Dunlosky, Tauber, & Rhodes (2008, Experiments 1-3)	Item	Intentional
Mueller, Dunlosky, Tauber, & Rhodes (2008, Experiment 4)	Pre-Study	Intentional
Kornell, Rhodes, Castel, & Tauber (2011)	Item	Intentional
Susser, Mulligan, & Besken (2013, Experiment 1)	Item	Intentional
OTHER		
Rhodes & Castel (2009) - Loudness	Item	Intentional
Alban & Kelly (2013) – Weight	Item	Intentional

Exploring the Potential Boundary Conditions

In the search for potential moderators of the disfluency effect, I created a grid organizing the literature among several dimensions. Tables 1 and 2 identify 21 studies employing various perceptual disfluency manipulations along with the central characteristics of the experiments described in the studies. Studies showing the disfluency effect are in Table 1. Studies not showing the disfluency effect are in Table 2. Together, Tables 1 and 2 list several potentially important moderating factors of the disfluency effect: type of disfluency manipulation, type of judgments of learning (JOLs), and encoding instructions. The tables are not exhaustive, nor do they depict every important moderating factor. I have included only studies that examine the disfluency effect in single-word experiments. Although there are many differences among the experiments in this collection, there also appear to be methodological practices that are associated with the failed replications and that provide a starting point for determining the boundary conditions of the disfluency effect.

Type of JOLs. One common feature in studies that did not obtain the disfluency effect is the use of item-by-item *judgments of learning* (JOLs). Such JOLs are predictions made by participants indicating their subjective belief (probability) that they will remember the studied words if tested at a later time. Researchers use JOLs to examine how students use various cues, such as fluency, to regulate the allocation of attentional resources during study and re-study (Metcalfe & Finn, 2008; Son & Metcalfe, 2000). JOLs can be elicited before studying (pre-JOLs; e.g., Muller, Dunlosky, Tauber, & Rhodes, 2014), immediately after studying each word (item-by-item JOLs; e.g., Rhodes & Castel, 2008), after the whole list is studied (aggregate or list-wide JOLs; e.g., Besken & Mulligan, 2014), or after a short delay (delayed JOLs; e.g., Dunlosky & Nelson, 1992). When the JOLs are made item-by-item, participants immediately judge their confidence on a scale of 0%-100%, with a JOL of 0% indicating that participants believe they will not be able to remember a word at a later time, and 100% indicating that they believe they will definitely recall the word at a later time As shown in Table 2, a majority of studies failing to show a memory benefit for disfluent items used item-by-item JOLs during encoding, while the majority of studies finding a disfluency effect did not. It is important to note that the primary focus in the studies employing item-by-item JOLs was on the impact of fluency on metamemory, with less emphasis on the disfluency effect. It is quite possible that the elicitation of JOLs may have had an impact on actual memory that might have masked a disfluency effect. For instance, item-by-item JOLs produce better memory for material compared to a condition that requires no item-by-item JOLs (e.g., Matvey, Dunlosky, Guttentag, 2001). Besken and Mulligan (2013) put forth two possible explanations as to why the elicitation of item-by item JOLs might attenuate the disfluency effect. One possible explanation is that elicitation of item-by-item JOLs requires the participant to retrieve the stimulus again, inducing deeper, more elaborative processing for *both* disfluent and fluent stimuli. Another possible explanation is that the elicitation of JOLs results in the allocation of more effort to the JOLs task, thereby leading to less post-lexical processing of the disfluent stimulus than would otherwise occur. The latter account is supported by the observation of participants' lower recall rates for disfluent items when giving item-by-item JOLs (Besken & Mulligan, Experiment 2) than when giving list-wide JOLs (Besken & Mulligan, Experiment 1). Regardless of which account is correct, the elicitation of item-by-item JOLs appears to influence processing and the subsequent memory strength of the memory representation. Thus, it is important to examine the influence of type of JOLs when examining disfluency.

Testing expectancy. Another potentially important moderating factor of the disfluency effect is encoding instructions. As seen in Tables1 and 2, there are differences in whether or not participants knew about an upcoming memory test. However, encoding instruction is confounded with the elicitation of JOLs. That is, when participants make a JOL, they most often are told there will be an upcoming memory test. In almost all cases presented in Tables 1 and 2, researchers used intentional encoding instructions in conjunction with JOLs. Thus, it is possible that encoding instruction, and not type of JOL, is the important factor. When individuals know about an upcoming memory test, they have better memory for the to-be-tested material compared to those who do not know about an upcoming memory test (e.g., Eitel & Khul, 2016; Van den Broek et al., 2001). Testing expectancy, then, may cause people to devote more resources to the encoding task, minimizing any processing differences obtained from perceptual disfluency. It is therefore important to examine the possible effect of testing expectancy on the disfluency effect, independent of the effect of item-by-item JOLs.

Type of disfluency manipulation. Another potential moderating variable is type of manipulation used to induce the disfluency effect. As can be seen in Table 1, several different disfluency manipulations have produced a disfluency effect: masking, inversion, auditory missing letters, and transposed letters. However, as seen in Table 2, several other perceptual fluency manipulations do not: font size color contrast, blurring, and font style. Who do some manipulations produce the disfluency effect while others do not?

While almost all types of perceptual disfluency incur processing costs (e.g., increased latencies), not all manipulations affect processing the same way. Manipulations can affect different loci (e.g. pre-lexical, lexical, and post-lexical) during processing. The locus impacted by a perceptual disfluency manipulation can be determined in a number of ways.

One way to determine the loci affected by a perceptual manipulation is via eye-tracking, which is a technique that provides a continuous measure of processing. The E-Z reader model (Reichele, Pollatsek, Fisher, & Rayner, 1998) was developed to provide a way to test the locus a manipulation influences in the context of natural reading (i.e., sentences), but the results can be generalized to single-word reading as well. Within the model, there are two stages of lexical identification that occurs during reading: An L_1 stage, which involves the encoding of featural and letter-level information, and an L_2 stage, which involves word and semantic-level information. Manipulations affecting the L_1 stage exert an effect on the word itself (*n*), while manipulations affecting the L_2 stage exert an effect on word n +1 (i.e., the word that immediately follows the target word). Identification of word $n + 1$ is tied to the completion of L_2 processing. That is, the longer it takes to complete the L_2 stage, the less parafoveal preview an upcoming word receives, resulting in increased fixation durations on word $n + 1$. Perceptual manipulations affecting an L_1 Stage are can be though to affect a pre-lexical locus, while manipulations exerting an influence on the L_2 stage can be said to affect a lexical or post-lexical locus.

Another way to determine the locus impacted by a perceptual disfluency manipulation is to use computational modeling (e.g., diffusion model; Gomez & Perea, 2014) or mathematical modeling (e.g., distributional analyses; Balota & Yap, 2011). Similar to eye-tracking, cognitive modeling allows researchers to more precisely examine the stages of information processing. One type of mathematical model commonly used is the ex-Gaussian distribution (see Balota & Yap, 2011 for a review). Parameters within the ex-Gaussian distribution are influenced by a variety of different experimental manipulations (e.g., word frequency, semantic priming, and stimulus degradation; Balota & Spieler, 1999). This mathematical model can be used to determine the locus or loci of a manipulation by decomposing the underlying RT distribution into three parameters: μ and σ (representing the mean and standard deviation of the Gaussian distribution) and τ (representing the exponential distribution or the skew of the distribution). These parameters appear to map on to different aspects of processing (but see Matzke & Wagenmakers, 2009, for a different view). For instance, a perceptual disfluency manipulation that shifts an entire reaction time (RT) distribution to the left or right on most or all trials indicates an early effect (what I am calling pre-lexical) and manifests as an effect on the μ parameters; a perceptual disfluency manipulation that affects the weight of the tail on some trials influences a late stage of processing (what I am calling lexical and post-lexical), which manifests on the τ parameter or all three parameters.

Using the techniques described above can help researchers identify the locus or loci of a perceptual disfluency manipulation. As noted before, two perceptual disfluency manipulations that have shown the disfluency effect are masking and word inversion. These two perceptual disfluency manipulations have been suggested to engender increased top down, higher-level, processing during word recognition (McClelland & Rumelhart, 1981; Wong, Twedt, Sheinberg, Gauthier, 2010). It is possible, then, that manipulations that affect only an early, pre-lexical, locus may not produce better memory. For example, in Hirshman et al. (1994, Experiment 5), luminance was used as a perceptual disfluency manipulation affecting a pre-lexical locus (i.e., words presented in dark-grey vs. white font on a black background). Lower luminance produced processing costs, denoted by longer naming latencies, but did not confer a memory benefit in free recall. Pre-lexical perceptual disfluency manipulations, like the one used in Hirshman et al., exert an additive influence on word recognition, with perceptual disfluency exerting an effect at the pre-lexical level (e.g., Gomez, Perea, Ratcliff, 2014; Reingold & Rayner, 2006; Sanchez & Jaeger, 2014; Yap & Balota, 2007). Thus, a manipulation affecting a pre-lexical stage of processing might produce no memory benefit (e.g., Hirshman et al., 1994; Yue et. al., 2012). Instead a manipulation affecting higher levels (i.e., lexical or post-lexical) might be needed to engender a disfluency effect. It is possible, then, that in order for the disfluency effect to confer any memory benefit, the perceptual disfluency manipulation must be strong enough to elicit increased feedback from lexical or post-lexical levels. An ideal way to test this would be to employ manipulations known to influence pre-lexical, lexical, and post-lexical levels of processing. This is more thoroughly examined in the experiments that follow.

Overview of Experiments

As previously noted, I complied a collection of 21 studies examining disfluency effect in single word memory. When the studies were organized according to central methodological characteristics, several potential moderating variables became apparent. Type of JOLs and encoding instruction seem to be the strongest moderators of the disfluency effect. Out of 21 studies, only two studies showed a disfluency effect with item-by-item JOLs and intentional instructions. Further, type of disfluency manipulation used by researchers also seems to be important.

In addition to the aforementioned variables, I examined three other potential moderators of the disfluency effect. The retention interval (RI) between study and test was examined. All the studies in Table 1 used short RIs (~3 min). If perceptual disfluency is to have any practical application, the disfluency effect must be demonstrated across longer intervals, similar to those most typically seen in classrooms. Only a few studies have shown the disfluency effect across longer RIs (e.g., 40 minutes to days or weeks; i.e., Diemand-Yauman et al., 2010; Weltman & Eakin, 2014), but these have been with longer materials (i.e., prose passages). Because other desirable difficulties, such as spacing and testing, tend to be more robust across longer RIs (e.g., Cepeda, Vul, Rohrer, Wixted, & Pashler, 2008; Roediger & Karpicke, 2006), it is important to show that the disfluency effect, if it can be found, can persist across longer delays.

Encoding duration and list type (i.e., whether words are presented in a blocked or mixed list) were also examined. These two variables are important for both theoretical and practical reasons. Encoding duration is important because pre-lexical and lexical accounts posit that encoding duration should have no impact on the disfluency effect whereas post-lexical accounts would predict otherwise (Hirshman et al., 1994). Experiment 4 examined these competing theoretical accounts. Finding that disfluency is moderated by whether words appear in a blocked or mixed fashion through varying list type would shed light on the mechanisms involved in the disfluency effect, because stronger disfluency effects in a mixed list would indicate support for distinctive processing models, whereas stronger disfluency effects in a blocked list would indicate support for a post-lexical locus for the effect. In addition, the question of whether or not the disfluency effect arises in mixed or blocked lists has practical (educational) implications. A disfluency effect that arises in only a mixed list would suggest using perceptual disfluency to enhance memory only rarely; in contrast, a disfluency effect found in both mixed and blocked lists calls for a more widespread use of perceptual disfluency as students' studying is almost always varied rather than blocked in nature.

Experiment 1 examined the influence of JOLs (item-by-item JOLs vs. aggregate JOLs) as a function of encoding instructions (incidental vs. intentional) on recognition memory while using a masking manipulation. Experiments 2, 3, and 4 explored the locus of the disfluency effect by examining two perceptual disfluency manipulations: low-level blurring and cursive handwriting and the role of encoding duration on the disfluency effect. Experiment 5 examined whether list design moderates the disfluency effect. Experiment 6 examined whether the benefits of disfluency extend over longer durations (e.g., 24 hours). Finally, all data were meta-analyzed to determine the overall effect size of the disfluency effect.

The goal of the current experiments is to further examine the disfluency effect by examining several boundary conditions. Perceptual disfluency manipulations may have potential pedagogical implications; thus it is important to establish the usefulness and robustness of this phenomenon. Overall, these experiments will help determine under what conditions using disfluency might promote positive learning outcomes in laboratory. This will provide a better understanding of the theoretical underpinnings of the disfluency effect.

CHAPTER 2. GENERAL METHOD AND PLAN OF ANALYSIS

All experiments followed the same general method and plan of analysis. The general method is described here along with the data analysis strategy, which includes statistical procedures and replacement criteria used. Unique aspects of each experiment are noted within their corresponding chapters.

Study and Test Phases

Materials. These studies used a selection of 200 nouns from the English Lexicon Project database (Balota et al., 2007) as stimuli (see Appendix A). I controlled for frequency (all stimuli were high-frequency; mean log HAL frequency = 9.2) and length (all stimuli were 4 letters in length). This is because both word frequency and length impact memory (e.g., Baddely, Thomson, & Buchana, 1975; Glanzer & Adams, 1990).

Apparatus. Stimuli were presented via E-Prime experimental software (Version 2.1; Schneider, Eschman, & Zuccolotto, 2002) on a Dell computer with an LCD monitor. A microphone attached to the E-prime SR response box recorded what participants said and how long it took participants to name the words. Participants used a keyboard to record their JOLs. Finally, the E-Prime SR response box was used to record the old/new recognition responses.

Procedure. Participants were tested individually in a small, well-lit room, seated approximately 65 cm from the computer screen. Before the experiment, all participants were informed that they would be naming words and that some of the stimuli would appear as hard-to-read (perceptually disfluent) words and some as easy-to-read (perceptually fluent) words. Participants engaged in four practice trials to familiarize themselves with the experimental procedure before the experiment proper. All fluent and disfluent words appeared in the center of the screen. Each trial began with a fixation cross appearing at the center of the screen for 1,000 ms (milliseconds). The fixation cross was used to ensure that participants' attention was directed to the location where the words appeared. The fixation cross was then replaced by a word in the same location. In order to ensure that all items were identified and to assess whether perceptually disfluent items were indeed disfluent, participants were instructed to name each word as quickly and accurately as possible. The naming latency and actual response were recorded on each trial. After the naming response, there was a 1,000 ms blank inter-stimulus interval (ISI). After all the words were named, participants who completed aggregate JOLs were told that fluent and disfluent words had appeared in the list, and were asked to estimate how many in each condition they expect to remember on a later test. The order of the memory judgments (fluent vs. disfluent) was counterbalanced across participants.

After the study phase, a short 3-minute distractor task was administered in which participants wrote down as many United States state capitals as they could. At the end of the three minutes, participants took an old-new recognition test. At test, a fixation cross appeared in the center of the screen for 1000 ms and was followed by a word that either had been presented during study ("old") or had not been presented during study ("new"). For each word presented, participants used the SR button box to input their responses. One button was labeled "old" to indicate that they had named the word during study, and another button was labeled "new" to indicate they did not remember naming the word. Words stayed on the screen until participants gave an "old" or "new" response. All words were individually randomized for each participant during both the study and test phases. After the experiment, participants were debriefed. Each entire experiment took a maximum of 30 minutes to complete.

Analysis

For all the experiments described, an alpha level of .05 is maintained. Cohen's *d* and generalized eta-squared (η_g^2; Olejnik & Algina, 2003) are used as effect size measures. Alongside traditional analyses that utilize null hypothesis significance testing (NHST), I also report the Bayes' factors for null findings when conducting my planned comparisons. One shortcoming of NHST is that it does not allow one to measure support that there is a true null difference between conditions. Within the NHST framework, a null result could occur if there truly is no effect of a manipulation, but also if the experiment did not have enough power. Additional Bayes' factor analyses is particularly important in the following set of experiments, as prior studies have failed to find a disfluency effect.

Bayes' factors allow one to provide evidence for a model relative to another model. Model 0 is the null hypothesis that there is no difference between conditions and Model 1 is the alternative hypothesis that there is a difference. All prior probabilities are Cauchy distributions centered at zero, and effect sizes are specified through the r-scale, which is the interquartile range (i.e., how spread out the middle 50% of the distribution is). For the null model, the prior is set to zero. The notation BF_{01} is used to express the probability of the data given the null hypothesis (Model 0) relative to the probability of the data given the alternate hypothesis (Model 1). Jefferys (1967) provides a general heuristic for interpreting Bayes factors. A Bayes' factor of 0-3 is weak evidence that there is a no true difference; 3-10 is substantial evidence for no true difference; 10-100 is strong evidence that there is no true difference; and > 100 is decisive evidence that there is no true difference. Bayes factors were calculated with the freeware software program JASP (Version 0.8.1.1; https://jasp-stats.org).

Naming accuracy and naming latency. In each experiment, items with more than 40% error rates across participants were discarded and not included in any of the analyses. Further, participants were required to have identified at least 75% of the items during study to be included in the experiment. If a participant did not meet this criterion, that participant was replaced. This was done to maintain a moderate to high level of accuracy in all reported experiments. Naming latencies faster than 150 ms or slower than 2.5 times the standard deviation for each participant were excluded from analysis as well.

JOLs. When item-by-item JOLs were obtained, the numeric entry was treated as a proportion. For aggregate JOLs, participant responses were transformed by dividing the number predicted to be recognized by the number of items present in each condition. The average proportion across participants for each condition (fluent and disfluent) was entered into the analysis.

Recognition memory. Performance was examined with *d'*, a memory sensitivity measure derived from signal detection theory (Macmillan & Creelman, 2004). Memory sensitivity (*d'*) is computed from the difference between the number of hits (i.e., judging an item as *old* when it is in fact *old*) and false alarms (i.e., judging an item *old* when it is *new*) and is a measure how well an individual can discriminate a stimulus from distractors. A larger *d'* value indicates an increased ability to discern a correct response from an incorrect response, while a smaller value indicates decreased ability to discern a correct or true response from an incorrect or false response. It is possible that a participant detects every signal, giving a hit rate of 1.00, or a participant might make no false alarms giving false alarm rate of 0. A right-tail *p*-value of 0 corresponds to z = infinity while a *p*-value of 1 corresponds to z = negative infinity. This poses a problem for computing *d'*. Thus, hits or false alarms of 0 or 1 were changed to .99 or .01.

CHAPTER 3. EXPERIMENT 1

Experiment 1 addressed the question of whether two post-lexical manipulations (the elicitation of JOLs during study and encoding instructions) influence the disfluency effect. As already described, JOLs (judgments of learning) are metamemory judgments participants are asked to make during study indicating how likely it is they will remember certain information on a later test. JOLs are used as a proxy for how individuals monitor and regulate their learning (Metcalfe & Finn, 2008; Son & Metcalfe, 2005). As shown in Table 1, however, experiments that use item-by-item JOLs often find no disfluency effect. The act of making a JOL for each individual word could mitigate the beneficial effects of perceptual disfluency on memory either by engendering more post-lexical processing of both fluent and disfluent stimuli or by taking away attention from the processing of disfluent stimuli (Besken & Mulligan, 2013).

Although the influence of perceptual fluency on JOLs and actual memory has been discussed previously (e.g., Besken & Mulligan, 2013, 2014), what is lacking in the current literature is a careful examination of the effect of JOLs and perceptual disfluency on memory in a single experiment. The current experiment used masking as the perceptual disfluency manipulation along with two different types of JOLs: item-by-item JOLs, which are elicited immediately after studying each word, and aggregate JOLs, which are elicited after all the items are studied. An aggregate JOLs procedure requires participants to estimate the total number of fluent and disfluent items they expect to remember at time of test. Aggregate JOLs should not have any influence on the strength of the memory representation, as JOLs are provided after all the words have been studied. In addition to examining whether type of JOLs moderates the disfluency effect, JOLs are used as a subjective measure of disfluency.

In addition to examining the impact of JOLs on disfluency, Experiment 1 examined the influence of encoding instructions (incidental vs. intentional). Although the requirement to make item-by-item JOLs may moderate the disfluency effect by inducing deeper processing of all words, the use of item-by-item JOLs also requires that participants receive intentional learning instructions; that is, individuals are told that they will have to remember the words for a later memory test. Aggregate JOLs can be obtained with either intentional or incidental instructions. It might be that having intentional instructions induces the deeper processing of all words rather than that deeper processing is the result of the item-by-item JOLs. Most of the studies showing a disfluency effect have used incidental instructions in that participants were not informed about the test until it occurred (e.g., French et al., 2014; Hirshman & Mulligan, 1991; Narine, 1988). Experiment 1 was designed to explicitly examine the role of encoding instruction (incidental versus intentional) without the confound of the presence/absence of making item-by-item JOLs.

This experiment compared three different groups: an item-by-item JOLs intentional instruction group, an aggregate JOLs intentional instruction group, and an aggregate JOLs incidental instruction group. The item-by-item group received instructions that alluded to a recognition memory test and were required to provide JOLs for each item during encoding. The aggregate JOLs intentional group received instructions that alluded to a recognition memory test and were required to provide aggregate JOLs after encoding. The aggregate JOLs incidental group were not told about a memory test but were required to provide aggregate JOLs after encoding. At study, naming latencies were examined to assess whether the masked words were in fact disfluent. At test, the disfluency effect was examined in a recognition memory task for each group, respectively.

To examine whether type of JOLs moderates the disfluency effect in recognition memory, the item-by-item JOLs intentional instruction group was compared to the aggregate JOLs intentional group. If item-by-item JOLs interfere with finding a disfluency effect, this would be evidence that type of JOLs moderates the disfluency effect. To examine whether testing expectancy moderates the disfluency effect in recognition memory, the aggregate JOLs incidental group was compared to the aggregate JOLs intentional group. If intentional instructions interfere with finding a disfluency effect, then this would be evidence that testing expectancy moderates the disfluency effect. If the disfluency effect arises because of deeper, more effortful, processing, item-by-item JOLs and testing expectancy should eliminate the effect, as both would countervail the effects of disfluency by eliciting additional processing for both fluent and disfluent stimuli (Eitel & Khul, 2016). Thus, finding that type of JOLs and/or testing expectancy moderate the disfluency effect would provide some evidence that the disfluency effect is impacted by post-lexical factors.

Method

Participants and Design

Eighty-four undergraduate students from Iowa State University participated for course credit; 28 students were assigned to each of the three groups. All participants were native speakers of English and with self-reported normal or corrected-to-normal vision. There were three between-subject groups: item-by-item JOLs with intentional instructions, aggregate JOLs with intentional instructions, and aggregate JOLs with incidental instructions. The within-subject variable was whether or not the words were masked.

Materials

Four counterbalancing lists were constructed from the 200 nouns described in Chapter 2. First, two separate 100-word lists were created, one to be used during study and test (old items) and one to be used only during test (new items). Next, two versions of each of these lists were created. Half the items were assigned to the perceptually disfluent condition and half to the perceptually fluent condition. Lists were assigned to participants so that across participants each word occurred equally often in the four possible conditions: masked old, nonmasked old, masked new, and nonmasked new. It is important to note that each new item was categorized as masked or nonmasked for counterbalancing purposes. All items on the test were presented without a mask.

Procedure

Before the experiment, all participants were informed that they would be naming words and that some words would appear as disfluent (i.e., masked) and some as fluent (i.e., unmasked). In addition, participants in the intentional encoding groups were told that they should try to remember the words presented for a later memory test. Each trial began with a fixation cross appearing at the center of the screen for 1,000 ms. The fixation cross was replaced by a word in the same location. Words were presented in a 44-point Courier New font in black on a white background. Half of the words were presented under masked conditions and half under nonmasked conditions. Masked words appeared for 80 ms and were backward masked for 1,920 ms; unmasked words appeared for 2,000 ms. After each naming response, in the aggregate JOLs groups, a blank 1,000 ms inter-stimulus interval (ISI) appeared. After all the words were named, participants were told that 50 fluent and 50 disfluent words had appeared in the list, and were asked to estimate how many in each condition they expect to remember on a later test. The order of the two memory judgments was counterbalanced across participants. In the intentional item-by-item JOLs group, immediately after naming each item, participants used the keyboard to rate their confidence on a scale of 0 (not confident at all) to 100 (very confident) that would be able to remember the item they studied 5 minutes from now.

Results and Discussion

No participants had to be replaced and no items were discarded.

Study Phase

To assess whether masked words were objectively disfluent, response latencies, error rates, and JOLs were examined with separate 3 x 2 mixed-model ANOVAs, with group (incidental aggregate JOLs, intentional aggregate JOLs, intentional item-by-item JOLs) included as a between-subjects variable and masking (masked vs. nonmask) included as a within-subjects variable The outlier procedure described in Chapter 2 resulted in the exclusion of 3% of the data from the incidental aggregate JOLs group, 3% from the incidental aggregate JOLs group, and 4% from the intentional item-by-item JOLs group. Trials in which there were microphone malfunctions (i.e., the microphone did not record a response) were also excluded (9% from the incidental aggregate JOLs group, 6% from the intentional aggregate JOLs group, and 10% from the intentional item-by-item JOLs group).

Naming accuracy and naming latency. Mean naming latencies and accuracy for each group are displayed in Table 3. Examining accuracy, the ANOVA indicated a main effect of masking, $F(1, 81) = 27.13$, $p < .001$, $\eta_g^2 = .11$. Participants were more error prone when naming masked words compared to nonmasked words. There was no effect of group or an interaction between masking and group, both $Fs < .78$, both $ps > .35$, both $BFs_{01} > 3$.

Examining latency, the ANOVA indicated no effect of masking, $F(1, 81) = 1.43$, $p = .236$, $\eta_g^2 = .00$, $BF_{01} = 3.12$. There was no interaction between masking and group, $F(2, 81) = .265$, $p = .768$. $\eta_g^2 = .00$, $BF_{01} = 5.00$. There was a main effect of group, $F(2, 81) = 18.47$, $p < .001$, $\eta_g^2 = .31$. Independent samples *t*-tests (Bonferroni-corrected) revealed that individuals took longer to name words in the item-by-item JOLs group compared to both the incidental aggregate group, $t = 5.25$, $p < .001$, $d = .57$ and the intentional aggregate JOLs group, $t = 5.28$, $p < .001$. $d = .58$. The longer latencies for the intentional item-by-item group might have arisen due to a task-switching decrement imposed by the item-by-item JOLs task. Overall, it appears that masking did not impose a slowdown in naming. Overall, while masking did not influence naming latencies, participants had a harder time naming masked words compared to nonmasked words, as revealed by mean accuracy rates.

Table 3

Mean Naming Accuracy (in proportions), Naming Latencies (in milliseconds), and JOLs (in proportions) for Words in Experiment 1 as a Function of Masking and Group

Condition	Naming Accuracy	Naming Latency	JOLs
Incidental aggregate JOLs			
Nonmasked	.99 (.00)	589 (16)	.57 (.05)
Masked	.98 (.01)	592 (22)	.50 (.04)
Difference	.01	3	.07
Intentional aggregate JOLs			
Nonmasked	.99 (.00)	592 (16)	.49 (.01)
Masked	.97 (.01)	577 (15)	.46 (.00)
Difference	.02	-15*	.03
Intentional item-by-item JOLs			
Nonmasked	.99 (.00)	811(43)	.54 (.04)
Masked	.98 (.01)	805(47)	.46 (.04)
Difference	.01	-6	.08

Note. Standard errors are shown in parentheses. * p < .05; ** p < .01; *** p < .001. Negative numbers signify that masked words were responded to faster than nonmasked words.

JOLs. Two participants had to be removed from the intentional aggregate JOLs analysis as a result of not providing a JOL judgment for each condition. Mean JOLs for each group are displayed in Table 3. The ANOVA indicated a main effect of masking, $F(1, 79) = 20.58$, $p < .001$, $\eta_g^2 = .02$. Participants predicted that they would recognize fewer masked words than nonmasked words. There was no effect of group and no interaction between masking and group, both *Fs* < 1.77, both *ps* > .17, both $BF_{01} > 3$. Using a subjective measure of disfluency, participants perceived masked words as being more disfluent than nonmasked words.

Overall, although participants were not slower in naming masked words, they were more error prone (had lower accuracy) and gave lower JOLs than to nonmasked words. This provides evidence that the masking manipulation was in fact disfluent.

Test Phase

Given the very high naming accuracy rates for both masked and nonmasked conditions in all three of the groups, I followed the recommended practice (Hirshman & Mulligan, 1991) of analyzing unconditionalized data (i.e., words named incorrectly at study have not been removed). Memory sensitivity (*d'*) for each group is displayed in Figure 1. The same 2 x 3 mixed-model ANOVA described earlier was used. The ANOVA indicated no main effect of masking, $F(1,81) = .869$, $p = .35$, $\eta_g^2 = .00$, $BF_{01} = 4.11$, and no interaction between masking and group, $F(2, 81) = 1.58$, $p = .212$, $\eta_g^2 = .00$, $BF_{01} = 1.14$. There was a main effect of group, $F(2,81) = 6.14$, $p = .003$, $\eta_g^2 = .12$. Prior research (e.g., Besken & Mulligan, 2012, 2013) has suggested that item-by-item JOLs may induce more effortful processing compared to aggregate JOLs. Indeed, a planned comparison (Bonferroni-corrected) revealed that memory tended to be better overall for the intentional-item-by-item group (*d'* = 1.93) compared to the intentional-aggregate group (*d'* = 1.65), $t(54) = 1.71$, $p = .09$, $d = .46$. This suggests that the elicitation of item-by-item JOLs produces better memory than aggregate JOLs due to increased post-lexical processing. Another planned comparison examined overall memory performance between aggregate-incidental and aggregate-intentional groups. It has been shown that intentional instructions engender deeper processing compared to incidental instructions (Eitel & Kuhl, 2016). Corroborating this, memory tended to better when participants were told about a memory test, $t(54) = 1.92$, $p = .06$, $d = .51$.

Although the interaction was not significant, planned comparisons were conducted examining the disfluency effect in each of the groups independently. Both type of JOLs and type of encoding instructions were examined as possible moderators of the disfluency effect. Type of JOLs were examined first by comparing the intentional item-by-item JOLs group to the intentional aggregate JOLs group. It was hypothesized that item-by-item JOLs, but not aggregate JOLs should eradicate the disfluency effect, as item-by-item JOLs induce more effortful processing during study than aggregate JOLs. Type of JOLs did not moderate the disfluency effect. That is, there was no disfluency effect in either the intentional item-by-item JOLs group, $t(27) = .56$, $p = .58$, $d = .10$, $BF_{01} = 9.59$, or the intentional aggregate JOLs group, $t(27) = .83$, $p = .42$, $d = .15$, $BF_{01} = 3.71$. This finding differs from other studies that did find that type of JOLs moderates the disfluency effect (e.g., Besken & Mulligan, 2013). Although type of JOLs did not moderate the disfluency effect in this study, it is important to note that in those studies testing expectancy was not examined as a potential moderating factor.

Next, testing expectancy was examined as a moderating factor of the disfluency effect. Eitel and Kuhl (2016) predicted that if the disfluency effect arises as a result of deeper, more effortful, processing, telling participants about a memory test should eliminate the effect, as testing expectancy would countervail the effects of disfluency by eliciting additional processing of fluent stimuli. Corroborating this, a disfluency effect was found in the incidental aggregate

JOLs group, $t(27) = 2.22$, $p = .04$, $d = .42$, but, as already noted, not in the intentional JOLs group. Thus, it appears that high testing expectancy moderates the disfluency effect and the moderation is likely via increased post-lexical processing when the test is expected.

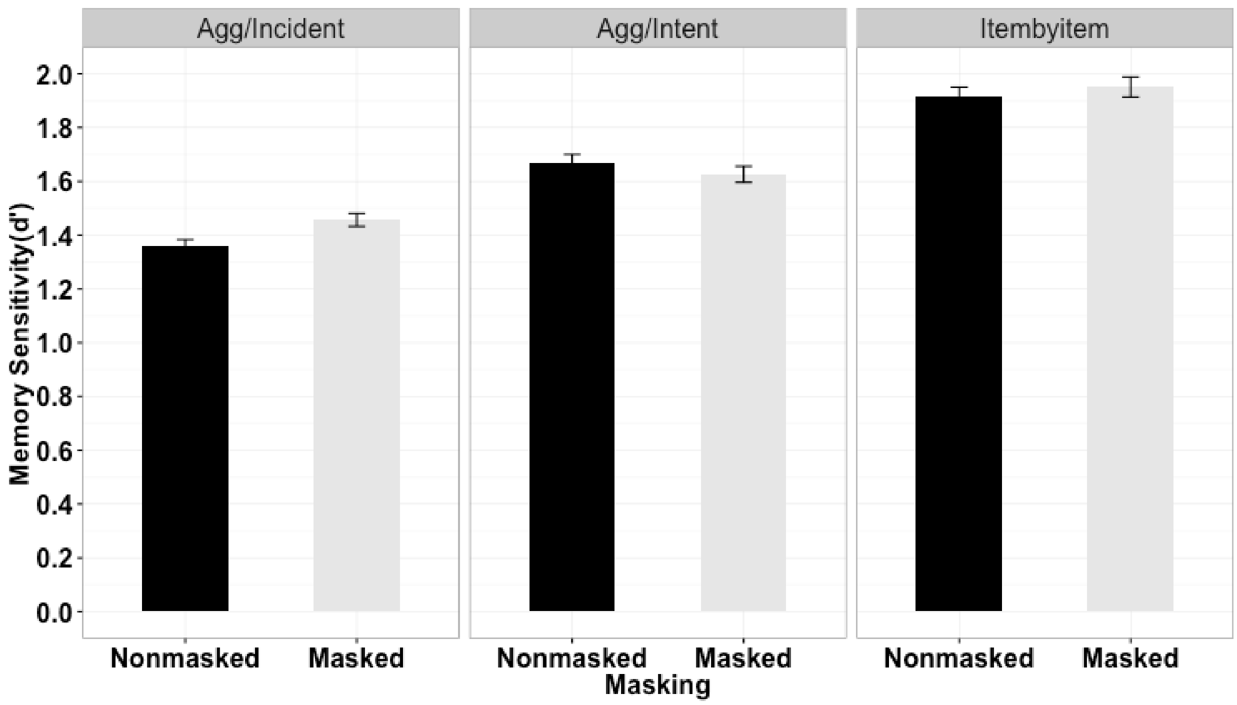

Figure 2. Memory sensitivity (*d'*) as a function of group (left - aggregate incidental group; center - aggregate intentional group; right - item-by-item intentional group (right)) and fluency (masked vs. masked). Error bars reflect the within-subject standard error of the mean (Morey, 2008).

Finding that testing expectancy moderates the disfluency effect supports post-lexical accounts of the disfluency effect (e.g., Atler, 2013; Eitel & Kuhl, 2016). Post-lexical accounts of the disfluency effect posit that better memory for disfluent items arises because of deeper, more effortful processing, after the stimulus has been recognized. It logically follows, then, that anything that prompts deeper processing should disrupt or eliminate the disfluency effect in learning. In the current study, telling students about an upcoming immediate memory test served to countervail the effects of disfluency—*both* fluent and disfluent stimuli received deeper processing, thus eradicating any beneficial effects of disfluency.

Overall, these findings provide evidence that a variable presumed to influence post-lexical processing (i.e., testing expectancy) can moderate the disfluency effect. The disfluency effect can be obtained as long as incidental instructions are used along with aggregate JOLs. These are the conditions under which the memory representations resulting directly from word processing can be examined without the impact of intentional learning effects. The next three experiments systematically examine whether the type of perceptual disfluency manipulation is an important moderator under these conditions. The role of a pre-lexical perceptual manipulation is examined next.

CHAPTER 4. EXPERIMENT 2

Experiment 1 found that encoding instructions moderate of the disfluency effect. That is, a disfluency effect was only found with incidental instructions. As described in Chapter 1, the disfluency effect has been attributed to processes arising at one of three loci: pre-lexical, lexical, or post-lexical. All perceptual disfluency manipulations will affect pre-lexical processing, but they may vary in their impact on the lexical and post-lexical loci. If the disfluency effect arises at a pre-lexical locus, it would reflect additional processing needed to resolve letter identity, but not word identity. A perceptual disfluency manipulation exerting an effect only on a pre-lexical locus would slow down recognition, but would not interfere or modulate later loci involved in word processing. If, however, the disfluency effect arises at a lexical locus, it would reflect increased interaction between the word level and the letter level. A perceptual disfluency manipulation affecting a lexical locus would not only slow down word recognition, but also modulate word processing by increasing the amount of top-down activation a word receives in order to recognize it. Finally, according to extant post-lexical, metacognitive, accounts of the disfluency effect (e.g., Alter, 2013), the memory benefit arises as a result of deeper, semantic, processing occurring after word recognition. From this post-lexical perspective, perceptual disfluency should be desirable across various disfluency manipulations, independent of the loci evoked by them. Thus, any perceptual disfluency manipulation, regardless of locus, should elicit a disfluency effect, if it is perceived as disfluent. Experiment 2 examined whether a disfluency effect can arise from a pre-lexical locus.

One manipulation that has been suggested to tap a pre-lexical locus is a low-level blurring manipulation (see Figure 3). In a natural reading study, Reingold and Rayner (2007) investigated the impact of degrading visual acuity (i.e., fainting), which is similar to blurring, on early (pre-lexical) and late (lexical) eye movement indices. The critical comparison was between a target word in a sentence that was fainted or presented normally. Reingold and Rayner showed that fainting only affected the encoding of letters at a pre-lexical locus (L_1 stage) that resulted in much longer fixation durations. Fainting, however, did not affect processing of a post-target word (i.e., increased fixations), which would have been associated with increased lexical processing (L_2 stage). Given that the visual acuity manipulation did not affect the L_2 stage to a greater degree than normal print, these findings suggest that degrading visual acuity affects only an early, pre-lexical, level of processing.

In the context of the disfluency effect with single words, Yue et al. (2013) used blurring as a pre-lexical disfluency manipulation. There was no disfluency effect, however, when they had participants provide item-by-item JOLs and participants were told about the upcoming memory test. Experiment 1 demonstrated that testing expectancy inhibited the disfluency effect. It is possible that blurring could show a disfluency effect without item-by-item JOLs and high test expectancy. In Experiment 2, low-level blurring was used to examine whether a perceptual disfluency manipulation thought to operate at a pre-lexical locus can produce a disfluency effect under incidental instructions.

Method

Participants and Design

Twenty-eight undergraduate students from Iowa State University participated for course credit. All participants were native speakers of English and with self-reported normal or corrected-to-normal vision. Clarity (blurred vs. clear) was used as a within-subjects variable and aggregate JOLs were used with incidental instructions. There was no between-subjects variable.

Materials and Procedure

The materials and procedure of Experiment 2 were as reported in Chapter 2. Blurring was used as the type of perceptual disfluency manipulation. Figure 3 depicts the two types of stimuli used. Both clear and blurred words were presented in 44-point Courier New font in black on a white background at the center of the screen. Blurred words, however, were distorted using Adobe Photoshop to reduce the pixels in each letter by 10% as was done by Yue et al. (2013). Words at study were presented for 500 ms. Context was reinstated at test. That is, studied items that appeared as blurred or clear also appeared as blurred or clear on the recognition test. In Experiment 1 this could not be done due to the masking manipulation employed.

Figure 3. Example of a 10% blur word on the right and a clear word on the left.

Results and Discussion

No participants were replaced or items discarded.

Study Phase

The outlier procedure described in Chapter 2 resulted in the exclusion of 3% of the data. Trials in which participants made errors or there were microphone malfunctions (i.e., the microphone did not record a response) were also excluded (10%).

Naming accuracy and naming latency. Mean naming accuracy and naming latencies appear in Table 4. Two paired samples *t* tests were used to examine naming accuracy and naming latency. Examination of naming accuracy showed no difference between clear and blurred words, $t(27) = .94$, $p = .35$, $d = .18$, $BF_{01} = 3.33$. Examination of naming latency was analyzed with a paired samples *t* test, which revealed that participants naming responses were 20 ms slower for blurred words compared to clear words, $t(27) = 3.72$, $p < .001$, $d = .70$. Using objective measures of disfluency, participants took longer to name blurred words compared to clear words, but did not make more errors.

Table 4

Mean Naming Accuracy (in proportions), Naming Latencies (in milliseconds) for Words in Experiment 2 as a Function of Clarity

Clarity	Naming Accuracy	Naming Latency	JOLs
Clear	1.00 (.00)	528 (20)	.54 (.05)
Blur	0.99 (.00)	548 (18)	.44 (.04)
Difference	.01	10	.10

Note. Standard errors are shown in parentheses.

JOLs. Mean aggregate JOLs are shown in Table 4. Participants predicted that they would recognize fewer blurred words than clear words, $t(27) = 3.11$, $p = .004$, $d = .59$. This suggests that blurred words were subjectively disfluent.

Taken together, both objective and subjective measures of disfluency showed that blurring made words perceptually disfluent.

Test Phase

Memory sensitivity (*d'*) is displayed in Figure 4. There was no indication that blurred words were better remembered than clear words, $t(27) = .07$, $p = .95$, $d = .01$, $BF_{01} = 4.98$.

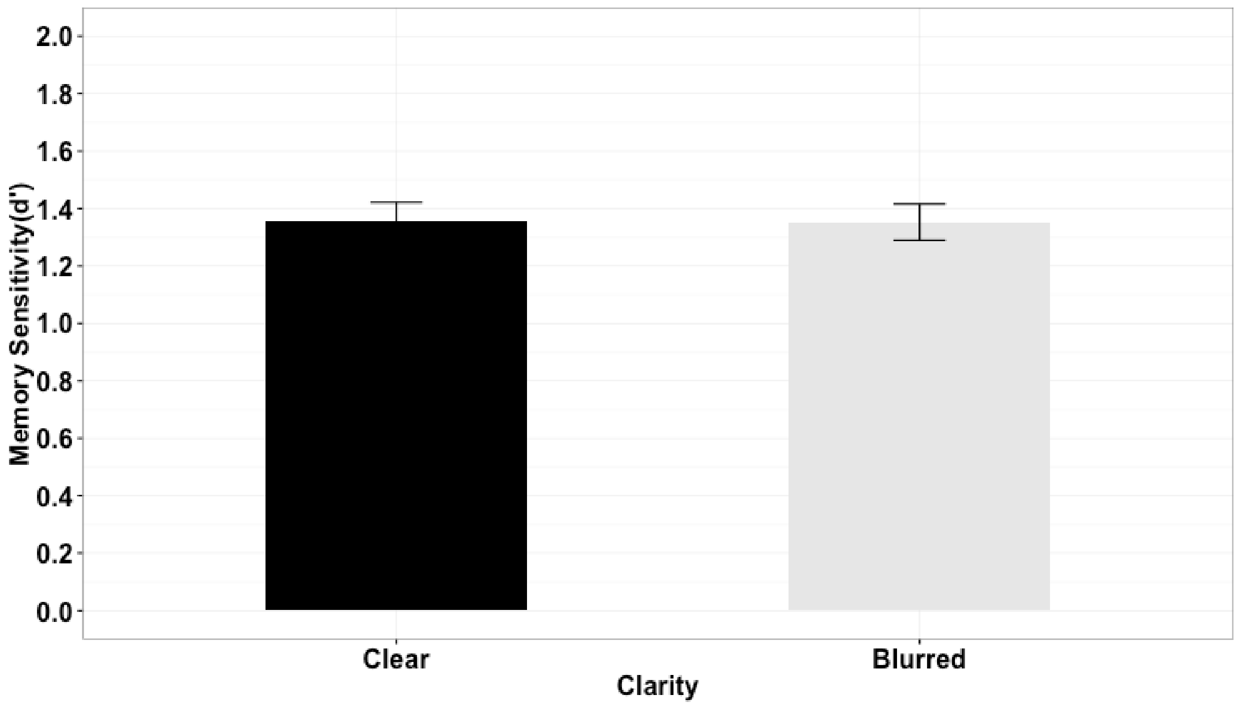

Figure 4. Memory sensitivity (*d'*) as a function of Fluency. Error bars reflect the within-subject standard error of the mean (Morey, 2008).

Experiment 2 examined whether a low-level blurring manipulation produces a disfluency effect. Yue et al. (2013) examined this with a blurring manipulation across several experiments and found no disfluency effect. In fact, Yue et al. found a tendency for blurred words to produce poorer memory than clear words. However, an important difference between their study and Experiment 2 was that they used item-by-item JOLs and intentional instructions, and as shown in

Experiment 1 the use of intentional instructions can affect how deeply stimuli are processed. In Experiment 2, item-by-item JOLs were removed and incidental instructions were used to more cleanly examine the effect of blurring on memory.

The blurring manipulation slowed down naming and participants' JOLs indicated that participants felt they would recognize fewer blurred words. However, blurring did not produce a disfluency effect. This finding is in accordance with Yue et al. (2013; but also see Hirshman et al., 1994, Experiment 5). If the disfluency effect were purely pre-lexical in nature, blurred words should have been remembered better than clear words. Taken together, these results argue against a purely pre-lexical locus. Moreover, these results also argue against a purely post-lexical locus account. A post-lexical account would predict a disfluent effect for blurred words as they were perceived as disfluent (i.e., lower JOLs). Thus, the disfluency effect might depend on a perceptual manipulation tapping a lexical locus of processing. This is examined in Experiment 3.

CHAPTER 5. EXPERIMENT 3

Experiment 2 examined whether blurring, which is presumed to operate only at a pre-lexical locus (Reingold & Rayner, 2007), produces a disfluency effect. There was no evidence that blurring produces better memory. What about perceptual manipulations that affect a lexical locus? Masking and inversion are two manipulations that have produced a disfluency effect (e.g., Hirshman & Mulligan, 1991; Nairne, 1988). Studies have suggested that those two manipulations make word recognition difficult enough that there is increased top-down activation from the word level to the letter level (McClelland & Rumelhart, 1981; Wong, Twedt, Sheinberg, Gauthier, 2010). Another perceptual disfluency manipulation that purportedly taps the same type of processing is handwritten cursive (Barnhart & Goldinger, 2010). Experiment 3 examined whether cursive handwriting produces a disfluency effect.

In an educational context, it is quite common for instructors to present students with cursive written information, either on the chalkboard/whiteboard or on the projector. Furthermore, it is not uncommon for students to write down notes from a lecture and study from them. Experiment 3 compared memory for words studied in handwritten cursive versus type print. If handwriting does produce a disfluency effect, cursive could be used to develop ecologically valid pedagogical interventions.

Handwritten stimuli differ in two keys ways from type-print stimuli: they are more ambiguous and they are non-segmented. In cursive writing, the same nominal characters change their physical forms across contexts; very similar forms may even signal different intended characters in different contexts. When individuals write in cursive, their letters connect together, creating potential problems in letter segmentation. These perceptual features create considerable difficulty in recognizing handwritten words. This difficulty is evinced by the fact that there is recruitment of frontal lobe areas responsible for attention when attending to handwritten stimuli (Qiao et al., 2010), as well as increased response latencies to handwritten stimuli as compared to type-printed stimuli (Barnhart & Goldinger 2010, 2015; Manso De Zunga et al., 1991; Perea, Gil-López, Beléndez, & Carreiras, 2016).

The increased difficulty associated with recognizing a word in cursive relative to type-print is thought to evoke increased top-down, lexical, processing. Indeed, Manso De Zuniga et al. (1991, Experiments 3 and 4) and Barnhart and Goldinger (2010, 2015) provided evidence for the increased lexical contribution by showing that certain lexical effects (word frequency, imageability, regularity, bidirectional consistency, and orthographic neighborhood size) are magnified during recognition of handwritten cursive words compared to recognition of type-printed words. Using an ex-Gaussian distribution (mentioned in the introduction), Barnhart and Goldinger (2010) found additional evidence that handwritten cursive magnifies lexical effects. Specifically, they found that script type (handwritten cursive vs. typed-print words) and word frequency (low vs high) interacted on both the μ and τ parameters, which are presumed to measure later processing, at a lexical or post-lexical loci. An effect on μ and σ, would reflect an early, pre-lexical, locus. This result provides evidence that handwritten word recognition is the result of increased interactivity between word and letter levels during the course of word recognition—that is, cursive evokes a lexical locus.

In Experiment 3, both easy-to-read and hard-to-read cursive were used to examine whether a stronger perceptual disfluency manipulation is needed to produce a desirable effect on memory. As already noted, cursive is purported to require more lexical processing in that recognition of a cursive word is purported to recruit increased word-level feedback (Barnhart & Goldinger, 2010). However, Perea et al. (2016a, 2016b) further suggested that the recognition of handwritten words is moderated by legibility of the cursive, with hard-to-read cursive recruiting more lexical feedback during word recognition than easy-to-read cursive. Examining RTs using quantile analyses (which complement an ex-Guassian analysis), Perea et al., (2016a) observed not only shifts in the RT distribution, but also skewing (i.e., a greater magnitude in higher quantiles) and more errors in the slower condition for both easy-to-read and hard-to read cursive words. However, hard-to-read cursive affected the lexical locus more; the word frequency effect was larger in magnitude for hard-to-cursive words. Additional support for the difference between easy-to-read and hard-to-read cursive comes from an eye-tracking study by Perea et al. (2016b). They found a substantial reading cost (i.e., more fixations, longer fixations, and longer reading times) for sentences written in cursive compared to type-print. At the word level, Perea et al. found that hard-to-read cursive words (compared to easy-to-read cursive) more strongly affected late measures of eye movements (e.g., total reading time and go past times) which are presumed to reflect lexical and post-lexical processes. If the disfluency effect reflects increased lexical processing as supposed by lexical accounts (i.e., compensatory processing account; Mulligan, 1996), it is possible that the memory benefit for processing words written in hard-to-read cursive will be larger than processing words written in easy-to-read cursive.

Experiment 3 examined whether cursive words (easy-to-read and hard-to-read) can induce a disfluency effect. At study, cursive words are expected to induce longer latencies, poorer accuracy, and lower JOLs. At test, accounts that place the locus at a lexical and/or post-lexical loci would predict an overall disfluency effect for cursive words, but different pattern of results are predicted for easy-to-read and hard-to-read cursive words. Accounts positing a lexical locus (e.g., the compensatory processing account; Mulligan, 1996) would predict a larger memory effect for hard-to-read cursive compared to easy-to-read cursive. This is because hard-to-read cursive not only slows down reading, but it also elicits increased interactivity between word and letter levels (Perea et al., 2016a, 2016b). It is this increased interactivity, according to the compensatory processing account, that produces better memory for disfluent words (Mulligan 1996). In contrast, extant theories positing a post-lexical locus would predict that cursive in general should produce a disfluency effect with no predicted difference between the different types of cursive. From a post-lexical perspective, both types of cursive slow down word recognition making it more difficult and less fluent, and therefore should stimulate the same post-lexical processing, which in turn, leads to a similar beneficial effect on memory performance.

To summarize, there are two questions concerning the use of cursive writing as a perceptual disfluency. The first question is whether cursive writing produces a disfluency effect. The second question is whether the magnitude of an observed effect varies as a nature of the difficulty of the cursive writing. To examine these questions, three levels of word fluency were used in Experiment 3 (type-print, easy-to-read cursive, and hard-to-read cursive). Within the context of a one-way related measures ANOVA, the main effect of script type is followed up by a Helmert contrast analysis, with two orthogonal contrasts. The first contrast evaluates the mean of the two cursive conditions (easy-to-read and hard-to-read) against the type-print condition; the second contrast compares easy-to-read to hard-to-read cursive. These contrasts examined whether cursive in general produces a disfluency effect and whether easy-to-read cursive or hard-to-read cursive produces better memory.

Method

Participants and Design

Thirty undergraduate students from Iowa State University participated for course credit. All participants were native speakers of English and with self-reported normal or corrected-to-normal vision. In Experiment 3, script (type-print, easy-to-read cursive, and hard-to-read cursive) was used as a within-subjects variable and aggregate JOLs were used along incidental instructions. There was no between-subjects variable.

Materials and Procedure

The materials and procedure of Experiments 3 were similar to what was described in Chapter 2. Two types of cursive (easy-to-read and hard-to-read) were included along with type-print. Because there are three presentation conditions (easy-to-read cursive, hard-to-read cursive, and type-print), only 99 words were presented at study (33 in each condition) and 198 words were presented at test. Lists were assigned to participants so that across participants each word occurred equally often in the six possible conditions: type-print old, easy-to-read cursive old, hard-to-read cursive old, type-print new, easy-to-read cursive new, and hard-to-read cursive new. Thus, counterbalancing required six lists rather than four.

Stimulus testing. To create the cursive stimuli, I asked seven individuals to write five sentences in cursive to be judged for penmanship. Forty students (N = 40) who did not know the purpose of the experiment rated the level of penmanship of the seven individuals on a 1–7 Likert scale (1 being legible and 7 being illegible). Participants were presented sentences, one at a time, by each writer in a randomized order in E-Prime.

I chose the individuals with the best and worst penmanship (average score of 2.2 vs. 4.0, respectively) to write the stimuli for the experiment (see Perea et al., 2016a, for a similar procedure).

The hard-to-read and easy-to-read stimuli were created using a Livescribe digital pen. The writing instrument resembles a normal pen with the addition of a small camera protruding from the tip. The camera reads a fine dot pattern on special paper, generating a digital trace of each pen stroke. The stimulus images were matched as closely as possible to the size of the computer-generated words. Figure 5 depicts an example of what a hard-to-read and an easy-to-read cursive word might look like. All stimuli used for this experiment can be found on the Open Science Framework (https://osf.io/yhqz4/).

Figure 5. Examples of type-print (left), easy-to-read cursive (center) and hard-to-read cursive (right) used across experiments.

Results and Discussion

No participants were replaced due to low accuracy (< 75%). Ten items from all script conditions were removed from all analyses due to error rates greater than 40%.

For all analyses, the main effect of script type is followed up with two orthogonal contrasts. This answers two questions. First, can cursive produce a disfluency effect? Second, is the memory benefit larger for easy-to-read or hard-to-read cursive?

Study Phase

Naming accuracy and naming latency. The same outlier procedure described in Chapter 2 resulted in the exclusion of 3% of the data. Trials in which there were microphone malfunctions were also excluded (6%). Mean naming latencies and accuracy rates appear in Table 5. For naming accuracy, a one-way repeated measures ANOVA (Greenhouse-Geisser corrected) indicated a significant effect of script, $F(2, 58) = 87.15$, $p < .001$, $\eta_g^2 = .63$. Planned contrasts revealed that participants were less accurate naming cursive words than type-print words, $t = 3.76$, $p < .001$, $d = .69$. Further, hard-to-read cursive words were more likely to be named incorrectly compared to easy-to-read cursive words, $t = 9.59$, $p < .001$, $d = 1.75$.

For naming latency, there was a significant effect of script, $F(2, 58) = 113.4$, $p = .021$, $\eta_g^2 = .30$. Planned contrasts revealed that participants named cursive words more slowly than type-print words, $t = 2.39$, $p < .001$, $d = .44$. Further, participants named hard-to-read cursive words more slowly than easy-to-read cursive words, $t = 5.38$, $p < .001$, $d = .98$. Together, the naming accuracy and latency confirm that the cursive stimuli were perceptually disfluent, with hard-to-read cursive being most perceptually disfluent.

Table 5

Mean Naming Accuracy (in proportions), Naming Latencies (in milliseconds), and JOLs (in proportions) for Words in Experiment 3 as a Function of Script Type

Script	Naming Accuracy	Naming Latency	JOLs
Type-print	.99 (.00)	536 (14)	.65 (.06)
Easy-to-read	94 (.01)	645 (19)	.54(.05)
Hard-to-read	.79 (.02)	695 (20)	.37 (.04)

Note. Standard errors are shown in parentheses

JOLs. Two participants were removed from the JOLs analysis because they did not provide JOLs for each stimulus type. Mean aggregate JOLs are shown in Table 5. A one-way repeated measures ANOVA indicated a significant effect of script, $F(2, 54) = 43.44$, $p < .001$, $\eta_g^2 = .16$. Planned contrasts revealed that participants predicted that they would recognize fewer cursive words than type-print words, $t = 4.37$, $p < .001$, $d = .83$. Further, participants predicted they would remember fewer hard-to-read cursive words than easy-to-read cursive words, t = 3.78, $p < .001$, d = .71.

Results from the study phase were clear-cut. As predicted, individuals were slower and less accurate when naming cursive stimuli than type-print stimuli, with hard-to-read cursive stimuli showing the longest response latencies and being the least accurate. In this study, however, accuracy was much lower than anticipated for hard-to-read cursive words. This is most likely due to the difficult nature of naming those stimuli coupled with the 500 ms encoding duration used. Nevertheless, the latency and accuracy pattern of findings mirrors results obtained from Perea et al. (2016). Aggregate JOLs also showed that participants believed that they would recognize fewer cursive words than type-print words, with participants believing they have the worst memory for the hard-to-read cursive stimuli. Overall, the subjective and objective measures indicate that cursive words (easy-to-read and hard-to-read) are disfluent.

Test Phase

For this and all subsequent analyses, words that were responded to incorrectly during the study phase were excluded from the *d'* analysis. The change to analyses conditionalized on naming accuracy is because accuracy rates for the hard-to-read and easy-to-read cursive conditions were moderately lower than type-print. Memory sensitivity (*d*') as a function of script type appears in Figure 6. A one-way repeated measures ANOVA (with the Greenhouse-Geisser correction applied) indicated no effect of script type, $F(2, 52) = 1.73$, $p = .194$, $\eta_g^2 = .03$, BF_{01} = 2.49. Planned contrasts revealed no memory difference between cursive and type-print, $t = .51$, $p = .61$, $d = .09$, $BF_{01} = 3.53$.

Further, there was marginal trend for hard-to-read cursive to produce poorer memory than easy-to-read cursive, t = 1.74, p = .09, d = .30, BF_{01} = 1.35.

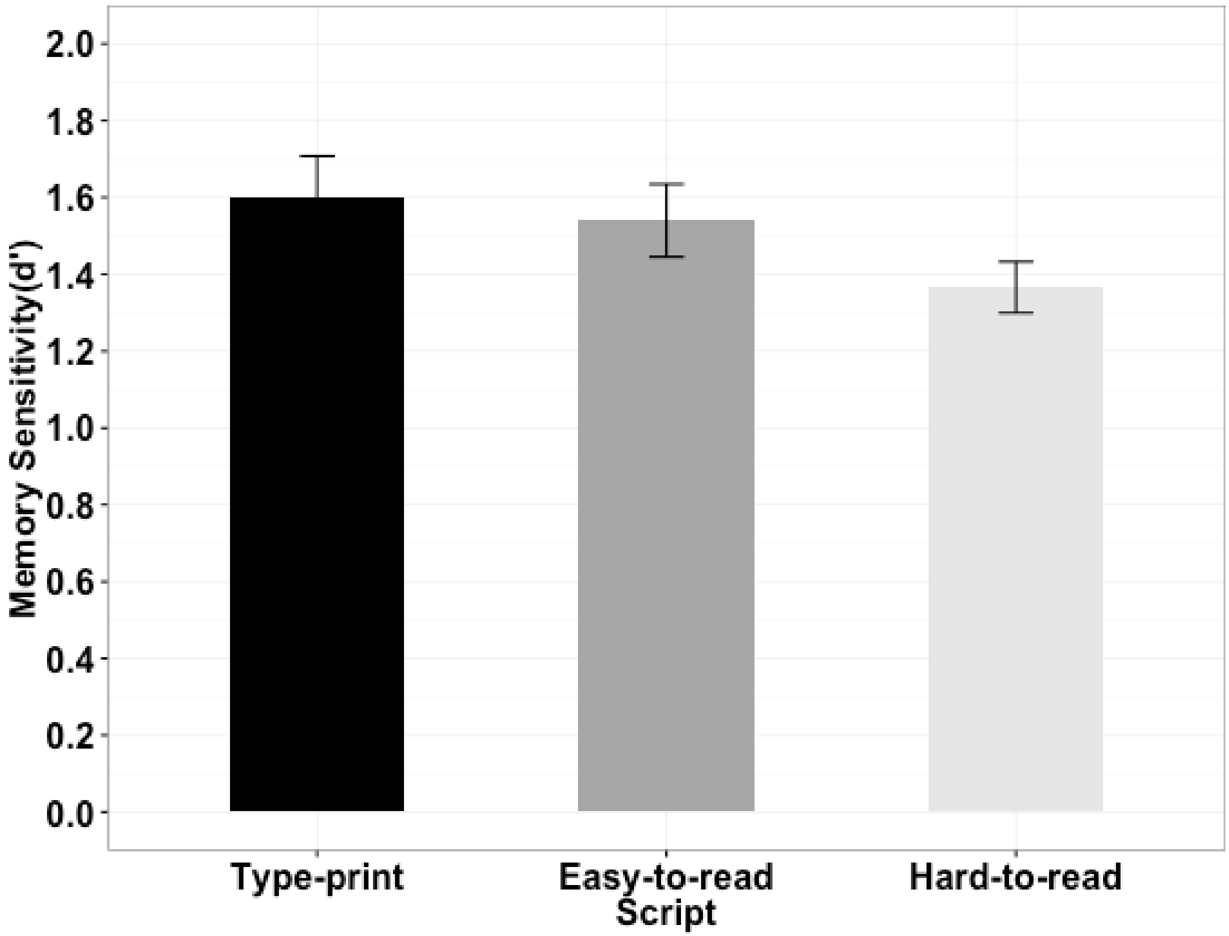

Figure 6. Memory sensitivity (*d'*) as a function of Script. Error bars reflect the within-subject standard error of the mean (Morey, 2008).

Experiment 3 examined whether the perceptual disfluency of cursive script could induce a disfluency effect. Contrary to all extant accounts of disfluency, no memory benefit arose for cursive words. A lexical account would predict a larger memory benefit for hard-to-read cursive words compared to easy-to-read cursive words. In contrast, participants had poorer memory for hard-to-read cursive stimuli than easy-to-read cursive stimuli. A post-lexical account would predict a cursive effect both easy-to-read and hard-to-read cursive stimuli. No benefit was found. Taken together, these results do not provide support for any of the extant disfluency theories.

In Experiment 3, word naming accuracy for the cursive words was much lower than for the disfluent words in the Experiments 1 and 2, especially for hard-to-read cursive. Perhaps the 500 ms presentation duration was insufficient to allow full processing of the cursive words. Additionally, there was some evidence that cursive hurts memory even when the word is recognized. Experiment 1 implicated a role for post-lexical processing in the disfluency effect. If lexical processing is incomplete, then post-lexical processing might also be inhibited. In order to test these possibilities, Experiment 4 was a replication of Experiment 3 with an increased encoding duration.

CHAPTER 6. EXPERIMENT 4

In Experiment 4, the task difficulty was lowered by increasing the study time for each word. The literature on the impact of study time on memory for disfluent stimuli is mixed. Yue et al. (2013, Experiments 2a and 3) found that extending encoding duration for blurred words from 2 s to 5 s brought memory performance up for blurred words to the level of clear words. In contrast, Hirshman et al. (1994, Experiment 1) found no evidence that increasing the encoding duration affected the magnitude of the disfluency effect obtained with masking.

In Experiment 4, the encoding duration was extended from 500 ms to 2 s to remove the possibility of incomplete processing during the study phase. Additional study time should allow more hard-to-read cursive items to be recognized and should allow for more post-lexical processing. Finding a disfluency effect for cursive stimuli when encoding duration is extended would provide support for a post-lexical account of the disfluency effect. If the results of the previous experiment are replicated, however, it would indicate that cursive processing has no beneficial effect on memory, and that hard-to-read cursive imposes too much demand on working memory, leading to superficial processing and poorer memory.

Method

Participants

Thirty undergraduate students from Iowa State University participated for course credit. All participants were native speakers of English and with self-reported normal or corrected-to-normal vision.

Materials, Procedure, and Design

The design of Experiment 4 was identical to that of Experiment 3, except that the encoding duration was increased from 500 ms to 2 s.

Results and Discussion

Two words were discarded from the analysis for errors greater than 40%. No participants were replaced.

Study Phase

Naming accuracy and naming latency. The outlier procedure resulted in the exclusion of 4 % of the data. Trials in which there were microphone malfunctions were also excluded (8%). Mean naming accuracy and latencies are shown in Table 6. Examining accuracy, a one-way repeated measures ANOVA indicated a main effect of script, $F(2, 52) = 56.35$, $p < .001$, $\eta_g^2 = .53$. Planned contrasts revealed that participants were less accurate when naming cursive words compared to type-print words, $t = 2.03$, $p = .047$, $d = .37$. Further, participants were also less accurate naming hard-to-read cursive compared to easy-to-read cursive words, $t = 10.42$, $p < .001$, $d = 1.90$.

Naming latency results indicated a significant effect of script, $F(2, 52) = 62.07$, $p < .001$, $\eta_g^2 = .34$. Planned contrasts revealed that participants tended to name cursive words more slowly than type-print words, $t = 1.92$, $p = .060$, $d = .35$. Further, hard-to-read cursive words were named more slowly than easy-to-read cursive words, $t = 4.32$, $p < .001$, $d = .79$. Together, the naming accuracy and latency confirm that the cursive stimuli were perceptually disfluent, with hard-to-read cursive being most perceptually disfluent.

Table 6

Mean Naming Accuracy (in proportions), Naming Latencies (in milliseconds), and JOLs (in proportions) for Words in Experiment 4 as a Function of Script Type

Script	Naming Accuracy	Naming Latency	JOLs
Type-print	.99 (.00)	588 (17)	.76 (.05)
Easy-to-read	.94 (.01)	739 (26)	.73 (.04)
Hard-to-read	.85 (.02)	825 (30)	.54 (.04)

Note. Standard errors are shown in parentheses

JOLs. Two participants had to be removed because JOLs were not provided for each script. Mean aggregate JOLs are shown in Table 6. Results indicated a significant effect of script, $F(2, 54) = 16.07$, $p < .001$, $\eta_g^2 = .15$. Planned contrasts revealed that participants predicted that they would recognize more cursive words than type-print words, $t = 2.34$, $p = .023$, $d = .43$. Further, participants predicted that they would recognize fewer hard-to-read cursive words than easy-to-read cursive words, $t = 4.61$, $p < .001$, $d = .84$. Using a subjective measure of disfluency, JOLs showed that participants believed that they would remember fewer cursive words than type-print words, with participants believing they would remember fewer hard-to-read cursive words.

Performance at study indicated that individuals were slower and less accurate when naming cursive stimuli than the type-print stimuli, with hard-to-read cursive stimuli showing the longest response latencies and being the least accurate. Examining aggregate JOLs, participants believed that they would recognize fewer cursive words than type-print words. Participants also predicted that they would recognize fewer hard-to-read cursive words compared to easy-to-read cursive words. Taken together the results indicate that cursive stimuli are perceptually disfluent.

Test Phase

As in Experiment 3, the analysis was conditionalized on correct naming responses. Memory sensitivity (*d')* values are displayed in Figure 7 as a function of script type. A one-way repeated measures ANOVA indicated a significant effect of script type, $F(2, 58) = 4.55$, $p = .01$, $\eta_g^2 = .05$. With an extended encoding duration (i.e., 2 s), planned contrasts indicated better memory for cursive compared to type-print, $t = 2.53$, $p = .014$, $d = .46$. However, there was no difference between easy-to-read cursive and hard-to-read cursive words, $t = 1.37$, $p = .176$, $d = .25$, $BF_{01} = 1.83$.

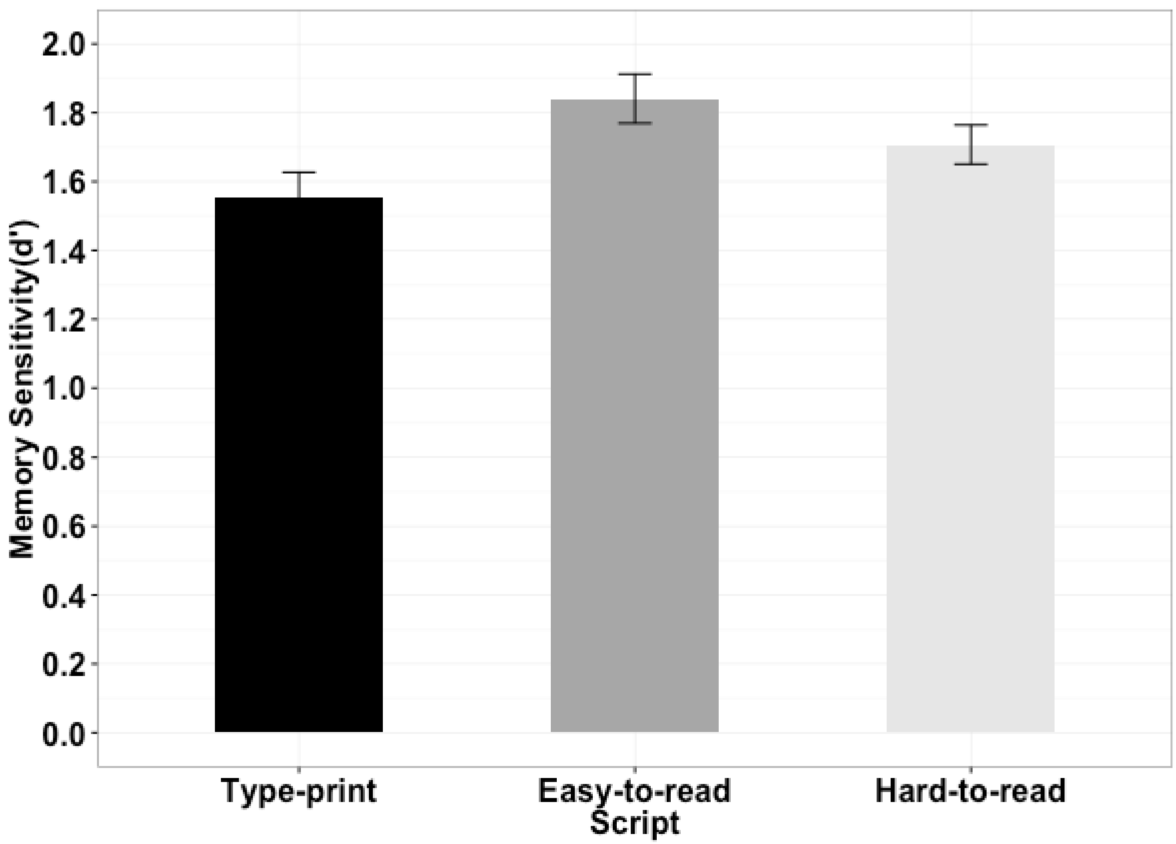

Figure 7. Memory sensitivity (*d'*) as a function of script with a 2 s encoding duration. Error bars reflect the within-subject standard error of the mean (Morey, 2008).

In Experiment 3, cursive stimuli did not produce a disfluency effect. In fact, presentation of hard-to-read cursive appeared to be an undesirable difficulty—hard-to-read cursive produced poorer memory compared to easy-to-read cursive. Further, participants, in general, had a difficult time reading hard-to-read cursive words. To examine if this was the result of difficulty associated with reading cursive, or the encoding duration employed, I increased the encoding duration to 2 s in Experiment 4. In contrast to Experiment 3 presentation of words in easy-to-read cursive and hard-to-read cursive produced a disfluency effect. Further, recognition memory was similar for easy-to-read and hard-to-read cursive words. The pattern of results observed suggests that while 500 ms (Experiment 3) is enough time to access the lexical representation for many cursive words, it is not enough time to deeply encode the cursive words. By increasing the study duration to 2 s, which presumably provides participants extra time to process stimuli, a very different pattern of results was obtained. That is, a disfluency effect was observed for both easy-to-read and hard-to-read cursive stimuli. This implicates a post-lexical locus to the disfluency effect.

This result is inconsistent with findings from Hirshman et al. (1994, Experiment 1) demonstrating that post-lexical processing does not modulate the disfluency effect. They presented a word for ~80 ms and manipulated how long the mask was presented after the word (i.e., 900 ms vs. 2500 ms). Increasing how long the mask was presented, and thus the amount of time that could be used for post-lexical processing of a stimulus, had no effect on the magnitude of the disfluency effect. They argued that this was evidence against a post-lexical locus. However, it is possible that in their study, masking an item for 900 ms allotted enough time to engage in sufficient post-lexical processing of identified words resulting in equivalent disfluency effects in the 900 ms mask and 2500 ms mask conditions. I have pilot data supporting this supposition. In one study, I presented words for~100 ms and quickly masked them for 400 ms. By presenting the mask for only 400 ms, I limited the amount of post-lexical processing that occurred. I did not find a disfluency effect when the mask was limited to 400 ms. Coupled with the fact that I did find a disfluency effect in Experiment 1 when the masked presented for 1920 ms, the current set of results provides evidence that a post-lexical locus does play a role in the disfluency effect.

Finding a disfluency effect of comparable magnitude for easy-to-read and hard-to-read cursive words is problematic for purely lexical accounts. The compensatory processing account would predict that hard-to-read cursive words should produce the best memory as the disfluency effect arises from the increased interaction between word and letter levels. In the current study, I did not find a recognition memory difference between easy-to-read cursive and hard-to-read cursive. However, there was a trend for recognition memory to be better for easy-to-read cursive, which would not be predicted by a lexical account. If the disfluency effect arises during word recognition, encoding duration should not influence the magnitude of the disfluency effect. In Experiment 3, presenting cursive words for 500 ms, while enough time to name them (results were conditionalized), did not result in better memory. However, increasing encoding time from 500 ms to 2 s did produce a disfluency effect. This indicates that the effect might be post-lexical in nature.

However, before making strong claims about the potential loci of the disfluency effect, it is important to establish the replicability of a disfluency effect produced by reading words in cursive. Experiment 5 examined list design (blocked vs. mixed) as another potential boundary condition of the disfluency effect.

The mixed list group is a close replication of Experiment 4. Manipulation of list design will illuminate the possible post-lexical mechanism operating to produce the disfluency effect.

CHAPTER 7. EXPERIMENT 5

Experiment 4 showed that cursive words can produce a disfluency effect when there is sufficient time for post-lexical processing. As noted in an earlier discussion in Chapter 1 on the theoretical mechanisms of the disfluency effect, previous research has implicated two post-lexical mechanisms that may underlie the disfluency effect—depth of processing and item distinctiveness. The depth-of-processing account would posit that cursive words produce a mnemonic benefit by receiving deeper processing as result of how perceptually disfluent, or hard-to-read, those items are. However, disfluency is confounded with distinctiveness (Diemand-Yauman et al., 2011; Rummer et al., 2016). That is, not only are cursive words disfluent, but they are also unusual. Given the rapid advances in use of technology in the modern world, fewer and fewer individuals are exposed to cursive writing, rendering this form of writing relatively unique.

Most studies examining the disfluency effect have used mixed designs in which a single list contains two levels of memoranda—a fluent level and disfluent level (e.g., Narnie, 1988; Hirshman et al., 1994; Mulligan, 1996; Sungkhasettee et al., 2011). In a mixed list, individuals may use a naïve (subjective) theory about what they believe requires more processing, and employ a strategy that involves differential attention to disfluent stimuli. This may induce participants to process disfluent stimuli more, thereby producing better memory. Indeed, a range of "desirable difficulties" (Bjork, 1994) show better memory in mixed-lists rather than blocked-lists. Examples include the production effect (e.g., Ozubko & Macleod, 2010), the generation effect (e.g., Kinoshita, 1989, Experiment 4), and even the testing effect (Mulligan, Susser, Smith, 2016). It may be that only when disfluent words occur in the background of fluent words during study that they stand out as distinctive and therefore receive the additional processing that produces differences in memory.

One way to test which post-lexical mechanism might be operating in the disfluency effect is by manipulating the design of the lists. If distinctiveness is removed as a cue by blocking the presentation of items, disfluency (distinctiveness) cannot be used to support differential processing, and the disfluency effect should be attenuated. By contrast, if a disfluency effect is still observed when the possibility of using distinctiveness as a cue is eliminated, this would be indicative of another mechanism, possibly depth of processing. It is possible that difficultly in recognizing cursive stimuli engenders deeper processing, thereby eliciting a better memory trace. Thus, manipulating list composition can be informative.

In Experiment 5, two groups are compared: a mixed-list group and a blocked-list group. It is hypothesized that, at study, cursive words will take longer to name, produce more errors, and be judged as more disfluent than type-print. At test, cursive words should show better memory than words studied in type-print. Finding this pattern would replicate the results of Experiment 4, and in fact the mixed list group closely replicates Experiment 4. Additionally, if list composition is indeed important, an interaction should be observed between script (easy-to-read, hard-to-read, and type-print) and list-type (mixed vs. pure). The presence of an interaction would indicate that distinctiveness may operate in producing the disfluency effect. This would suggest that individuals are using perceptual disfluency as a cue to direct differential attention. If there is no interaction, this would suggest that depth of processing is a likely candidate for producing the memory benefit underlying the disfluency effect.

Method

Participants

Seventy-two participants from Iowa State University participated for course credit. Thirty-six students were assigned to each of the two groups. All participants were native speakers of English and with self-reported normal or corrected-to-normal vision.

Materials, Procedure, and Design

The materials and procedure were an extension of Experiment 4. For the mixed-list group, cursive and type-print words were randomly intermixed by individual for the study and test phases. This replicated the procedure used in Experiment 4. For the blocked-list group, during the study phase, the order of presentation of items was random by participant, with the constraint that all type-print, easy-to-read cursive, and hard-to-read cursive were shown in separate blocks. Block order was counterbalanced across participants. Blocks were separated by a short break, in which the participants were told to take a break before proceeding on to the next block. For the blocked-list group, during the test phase, all items were intermixed randomly by individual. Thus, list type (blocked versus mixed) was manipulated as a between-subjects factor and script (type-print vs. easy-to-read cursive vs hard-to-read cursive) as a within-subjects variable.

Results and Discussion

Two words were discarded prior to all analyses reported herein due to error rates being greater than 40%. No participants had to be replaced due to low accuracy.

Study Phase

Naming accuracy and naming latencies. The outlier procedure detailed in Chapter 2 resulted in the exclusion of 4% of the data (2% for the blocked group and 2% from the mixed group). Trials in which there were microphone errors (11%) were also excluded from the latency data prior to analysis. Mean naming accuracy and latency are shown in Table 7. Naming accuracy and naming latencies were submitted to a 2 x 3 mixed-factor ANOVAs, with list type (blocked versus mixed) as a between-subjects factor and script (type-print vs. easy-to-read cursive vs hard-to-read cursive) as a within-subjects variable.

The accuracy analysis (Greenhouse-Geisser corrected) revealed a main effect for script, $F(2, 140) = 143.75$, $p < .001$, $\eta_g^2 = .52$. Planned contrasts showed that participants were less accurate naming cursive words compared to type-print words, $t = 4.20$ $p < .001$, $d = .50$. Further, individuals were also less accurate naming hard-to-read cursive stimuli compared to easy-to-read cursive stimuli, $t = 11.83$, $p < .001$, $d = 1.40$. The main effect for list type and the interaction between script and list type were not significant, both *Fs* < 1.99, both *ps* > .10, and $BF_{01} > 3$.

Table 7

Mean Naming Accuracy (in proportions), Naming Latencies (in milliseconds), and JOLs (in proportions) for Words in Experiment 5 as a Function of Script Type and Design

Script	Naming Accuracy	Naming Latency	JOLs
Mixed List			
Type-print	.99 (.00)	588 (17)	.78 (.04)
Easy-to-read	.94 (.01)	739 (26)	.71 (.04)
Hard-to-read	.85 (.02)	825 (30)	.51 (.04)
Blocked List			
Type-print	.99 (.00)	649 (17)	.83 (.04)
Easy-to-read	.93 (.01)	813 (25)	.75 (.04)
Hard-to-read	.81 (.02)	915 (28)	.57 (.04)

Note. Standard errors are shown in parentheses

The latency analysis revealed a significant effect for script type, $F(2, 140) = 206.97$, $p < .001$, $\eta_g^2 = .46$. Planned contrasts revealed that although the means were in the expected direction, cursive words were not named more slowly than type-print, $t = 1.57$, $p = .112$, $d = .19$, $BF_{01} = 2.40$. However, participants were slower to name hard-to-read cursive words compared to easy-to-read cursive words, $t = 20.69$, $p < .001$, $d = 2.44$. There was no main effect of list type, $F(1, 70) = .477$, $p = .49$, $\eta_g^2 = .00$, $BF_{01} = 4.37$. There was a marginal interaction between list type and script, $F(2, 140) = 2.56$, $p = .08$, $\eta_g^2 = .04$. This interaction reflected the fact that cursive words were not named more slowly in the blocked list, $t = 1.58$, $p = .119$, $d = .19$, $BF_{01} = 2.36$, but were marginally named more slowly in the mixed list, $t = 1.86$, $p = .067$, $d = .22$. In both groups, hard-to-read cursive words were named more slowly than easy-to-read cursive, both *ts* > 13.66, *ps* < .001. Together, the naming accuracy and latency confirm that the cursive stimuli were perceptually disfluent, with hard-to-read cursive being most perceptually disfluent.

JOLs. Data from one participant had to be removed from the JOLs analysis because JOLs were not provided for each script type. Mean aggregate JOLs are shown in Table 7. The JOLs analysis revealed a main effect of script, $F(2, 140) = 67.49$, $p < .001$, $\eta_g^2 = .17$. Planned contrasts revealed that participants predicted that they would recognize fewer cursive words compared to type-print words, $t = 3.13$, $p < .001$, $d = .37$. Further, participants said that would recognize fewer hard-to-read cursive compared to easy-to-read cursive words, $t = 8.13$, $p < .001$, $d = .96$. The main effect of list type and the interaction between list type and script were not significant, both *Fs* < .913, *ps* > .10, and $BFs_{01} > 3$. Overall, using a subjective measure of disfluency, JOLs showed that participants believed that they would recognize fewer cursive words than type-print words, with participants believing they would remember the fewest hard-to-read cursive words.

Performance at study indicated that individuals were less accurate when naming cursive stimuli than the type-print stimuli, with hard-to-read cursive stimuli showing the worst performance. Although naming latency did not show the expected cursive effect, response latencies were similar to those observed in previous experiments. There did not appear to be any speed-accuracy trade-off.

Examining aggregate JOLs, participants believed that they would recognize fewer cursive words than type-print words. Participants also predicted that they would recognize fewer hard-to-read cursive words compared to easy-to-read cursive words. Again, both objective and subjective measures of disfluency confirmed that the cursive words used were in fact disfluent.

Test Phase

An analysis of *d'* was conducted using the same 2 x 3 mixed-factor ANOVA described earlier. Similar to Experiments 3 and 4, the analysis was conducted on conditionalized data. Memory sensitivity (*d')* values are displayed in Figure 8. The analysis revealed a main effect of script, $F(2, 140) = 4.46$, $p = .013$, $\eta_g^2 = .02$. Planned contrasts revealed that recognition memory for cursive words was better than type-print, $t = 2.61$, $p = .010$, $d = .31$. There were no recognition memory differences between easy-to-read cursive and hard-to-read cursive words, $t = 1.53$, $p = .128$, $d = .18$, $BF_{01} = 2.19$. The main effect for list type and the interaction between script and list type were not significant, both $Fs < .263$, $ps > .10$, and $BFs_{01} > 3$. Once again cursive showed a disfluency effect, with no difference between easy-to-read and hard-to-read cursive. The cursive disfluency effect was similar across list types.

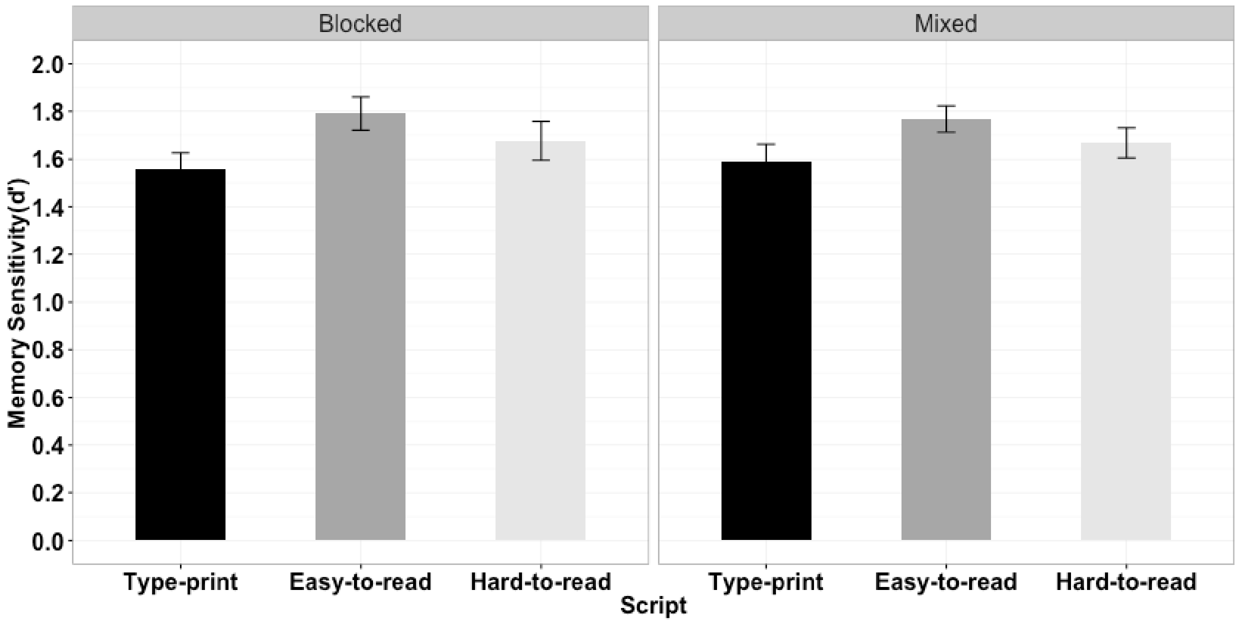

Figure 8. Memory sensitivity (*d'*) as a function of list type and script for Experiment 5. Error bars reflect the within-subject standard error of the mean (Morey, 2008).

An important result was that recognition memory was better overall for cursive stimuli. Further, there was no difference in memory between easy-to-read and hard-to-read cursive words. This pattern of findings replicates Experiment 4. Extant post-lexical accounts (e.g., Alter, 2013) would predict that the cursive words would show a disfluency effect, and that there should be no difference between the two types of cursive. Thus, this experiment provides further evidence that the disfluency effect arises at a post-lexical locus.

A second important result is that the disfluency effect was not affected by list design. Indeed, the pattern between groups was virtually the same (see Figure 8). The statistically indistinguishable disfluency effect for blocked and mixed lists suggests that distinctiveness is not the post-lexical mechanism driving the disfluency effect, rather depth of processing induced by the difficulty of processing the stimulus is the most likely mechanism. This is in accord with other studies demonstrating that distinctiveness plays only a small role in the disfluency effect

(Hirshman et al., 1994; Sungkasette et al., 2011). Further, these results suggest that the disfluency effect may have a different underlying mechanism from other desirable difficulties such as production, generation, and even masking (MacLeod et al., 2010; McDaniel & Bugg, 2008; Westerman & Greene, 1997), which are most robust in mixed designs. Future research should examine the differences between cursive and other perceptual disfluencies.

Experiments 4 and 5 showed that cursive words can act a desirable difficulty on an immediate recognition memory test. In order to be a true desirable difficulty, an effect must persist across longer retention intervals. Indeed, some desirable difficulties (e.g., spacing) become more robust after longer delays (Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006). Experiment 6 examined a longer retention interval (24 hours).

CHAPTER 8. EXPERIMENT 6

A major limitation in the research on perceptual disfluency is the use of short retention intervals (RIs). RIs are delays between study and when the final test is given. All the studies in Table 1 used short retention intervals to examine the disfluency effect (~3 min). There have been some studies that have used longer retention intervals (e.g., Diemand-Yaueman et al., 2010). These studies, however, use longer, more complex, study materials. The longest retention interval was used by Diemand-Yaueman et al. (2010) when they tested the effects of disfluency by putting PowerPoint presentations in a hard-to-read font. In their Experiment 2, the final test occurred weeks or months after being exposed to the material. In an educational context, the time lapse between study and test could be hours, days, or weeks. Other desirable difficulties that promote learning (e.g., testing and spacing) persist and become more robust across longer retention intervals, spanning weeks (Carpenter, Pashler, Wixted, & Vul, 2008) or months (Carpenter, Pashler, & Cepeda, 2009; Larsen, Butler, & Roediger, 2009). Thus, if perceptual disfluency is indeed a desirable difficulty it should persist over longer intervals. Experiment 6 examined the disfluency effect using a more educationally realistic 24-hour RI.

Method

Participants

Forty-two undergraduate students from Iowa State University participated for course credit. All participants were native speakers of English and with self-reported normal or corrected-to-normal vision.

Materials, Procedure, and Design

The materials and procedure were identical to Experiment 4, with one exception: the length of the RI was extended from 3 minutes to 24 hours. After the study phase, individuals were told to come back 24 hours later for the second part of the experiment. No mention was made of the recognition test (i.e., incidental instructions). Participants were not debriefed until the end of the second session. Script was manipulated as a within-subjects variable using a mixed-list design with incidental instructions.

Results and Discussion

Two participants were replaced due to accuracy less than 75%. Two words were also discarded due to error rates greater than 40%.

Study Phase

Naming accuracy and naming latency. The outlier procedure resulted in the exclusion of 4 % of the data. Trials in which there were microphone malfunctions were also excluded (7%). Naming accuracy and latency are shown in Table 8. Examining accuracy, the ANOVA (Greenhouse-Geisser corrected) indicated a main effect of script, $F(2, 82) = 96.72$, $\eta_g^2 = .52$. Planned contrasts revealed that participants were less accurate naming cursive words compared to type-print words, $t = 2.60$, $p = .011$, $d = .40$. Participants were less accurate naming hard-to-read cursive words than easy-to-read cursive words, $t = -8.78$, $p < .001$, $d = 1.35$.

One person was excluded from the latency analysis due to naming latencies not being recorded on the majority of the trials (>95%). Examining naming latency, the analysis revealed a significant effect of script, $F(2, 80) = 163.6$, $p < .001$, $\eta_g^2 = .39$. Planned contrasts revealed that participants took longer to name cursive words compared to type-print words, $t = 2.89$, $p = .005$,

d = .45. Further, participants named hard-to-read cursive words more slowly than easy-to-read cursive words, t = -6.42, p < .001, d = 1.00. Together, the naming accuracy and latency confirm that the cursive stimuli were perceptually disfluent, with hard-to-read cursive being most perceptually disfluent.

Table 8

Mean Naming Accuracy (in proportions) Response Latencies (in milliseconds), and JOLs (in proportions) for Words in Experiment 6 as a Function of Script Type

Script	Naming Accuracy	Naming Latency	JOLs
Type-print	.98 (.01)	634 (15	.77 (.04)
Easy-to-read	.92 (.02)	790 (20)	.74 (.04)
Hard-to-read	.80 (.02)	877 (23)	.55 (.04)

Note. Standard errors are shown in parentheses

JOLs. Mean JOLs are shown in Table 8. A one-way repeated measures ANOVA indicated a significant effect of script, $F(2, 82) = 36.21$, p < .001, $\eta_g^2 = .15$. Planned contrasts revealed that participants predicted that they would recognize fewer cursive words than type-printed words, t = 3.11, p = .003, d = .48. Participants predicted that they would recognize fewer hard-to-read cursive words compared to easy-to-read cursive words, t = 6.65, p < .001, d = 1.03. The analysis of JOLs paralleled that found in the objective data.

The results indicate that the cursive stimuli used in the current experiment were objectively and subjectively disfluent.

Test Phase

Similar to Experiments 3, 4 and 5, the analysis was conducted on conditionalized data. Memory sensitivity (d') for each script type is displayed in Figure 9. As would be expected with a 24-hour delay, memory was lower than in previous experiments. The ANOVA indicated a significant effect of script type, $F(2, 82) = 9.35$, $p < .001$, $\eta_g^2 = .09$. Planned contrasts showed that after a 24-hour delay, participants had better recognition memory for cursive than type-print words, $t = 3.55$, $p < .001$, $d = .55$. Further, participants tended to remember more easy-to-read cursive words than hard-to-read cursive words, $t = 1.84$, $p = .07$, $d = .28$.

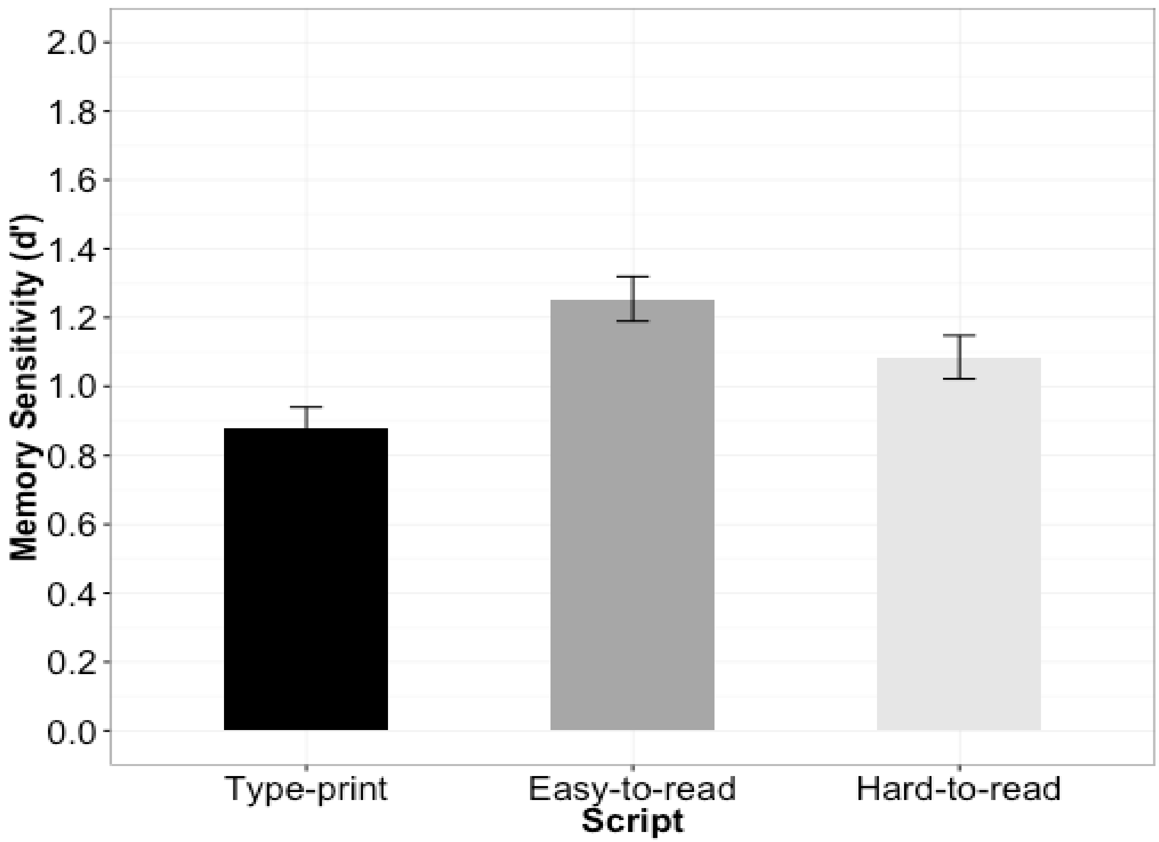

Figure 9. Memory sensitivity (*d'*) as a function of script after 24 hours. Error bars reflect the within-subject standard error of the mean (Morey, 2008).

The results from Experiment 6 are straightforward. Even after 24 hours, studying easy-to-read cursive and hard-to-read cursive had a desirable effect on memory, replicating the pattern in Experiments 4 and 5. However, easy-to-read cursive words were better remembered compared to hard-to-read cursive, a finding that was not observed in the previous experiments, although the trend was present. It is not uncommon for effects to become stronger or more robust with a longer delay. In fact, other "desirable difficulties," such as testing (e.g., Roediger & Karpicke,

2006) and spacing (e.g., Glenberg, 1976), become stronger with a delay. It appears the cursive disfluency effect persists over longer delays and that forgetting might be slower when the word is slightly disfluent (easy-to-read cursive) compared to highly disfluency (hard-to-read cursive). Overall, it does not appear that the RI moderates the disfluency effect as it has now been observed with an immediate test (Experiments 4 and 5) and a delayed test taken 24 hours later (Experiment 6).

CHAPTER 9. META-ANALYSIS OF EXPERIMENTS 4-6

In Experiments 4, 5, and 6 presenting words in cursive for 2 s served as a desirable difficulty. That is, cursive words were better remembered than type-print words. However, examining the critical comparison between easy-to-read and hard-to-read cursive words resulted in inconclusive findings. In Experiments 4 and 5 there was a trend for easy-to-read cursive words to be better remembered than hard-to-read cursive, but the effect did not achieve statistical significance. In Experiment 6, however, this difference was marginally significant. Given the inconsistent evidence and the theoretical importance of this comparison, a meta-analysis—a statistical technique used to synthesize results from multiple studies—was utilized to provide resolution as to the true nature of this effect. Because meta-analyses pool evidence from multiple studies, it is thought to be a more powerful indicator that a true effect is present (McShane & Bockenholt. 2017). Based on recent recommendations (e.g., Braver, Thoemmes, & Rosnthal, 2014; Cumming, 2013; Lakens & Etz, 2017; McShane & Bockenholt), I conducted a small-scale meta-analysis in R using the *metafor* package (Viechtbauer, 2010). In addition to examining the effect size between easy-to-read and hard-to-read cursive, an overall meta-analytic effect size is computed for the difference between easy-to-read cursive and type print and hard-to-read cursive and type-print. For each comparison, a random effects model integrated effect sizes across Experiments 4, 5, and 6 (k= 3, N = 144).

The average effect sizes and the confidence intervals (CIs) for each comparison of interest are reported in Figure 10. The meta-analysis revealed that the disfluency effect was present with easy-to-read cursive, *estimate* = .46, SE = .14, z = 3.94, $p < .001$, and hard-to-read cursive, *estimate* = .22, SE = .09, z = 2.67, $p < .01$. Further, easy-to-read cursive had a larger effect on memory than hard-to-read cursive stimuli, *estimate*= .23 *SE*= .08, $z = 2.75$, $p < .01$. The magnitude of the effects fluctuated across studies, however, given that they were within the CIs, the fluctuations are likely due to sampling error (Cumming, 2014). Indeed, the heterogeneity of each effect of interest was not statistically significant (see Figure 9). Overall, the meta-analysis showed a disfluency effect for both easy-to-read and hard-to-read cursive words. Further, the meta-analysis showed that the difference between easy-to-read and hard-to-read cursive words was small, but nonzero, with easy-to-read cursive words producing better memory. The implications of these and the rest of the findings as it relates to the potential mechanisms underlying the disfluency effect are discussed in the general discussion.

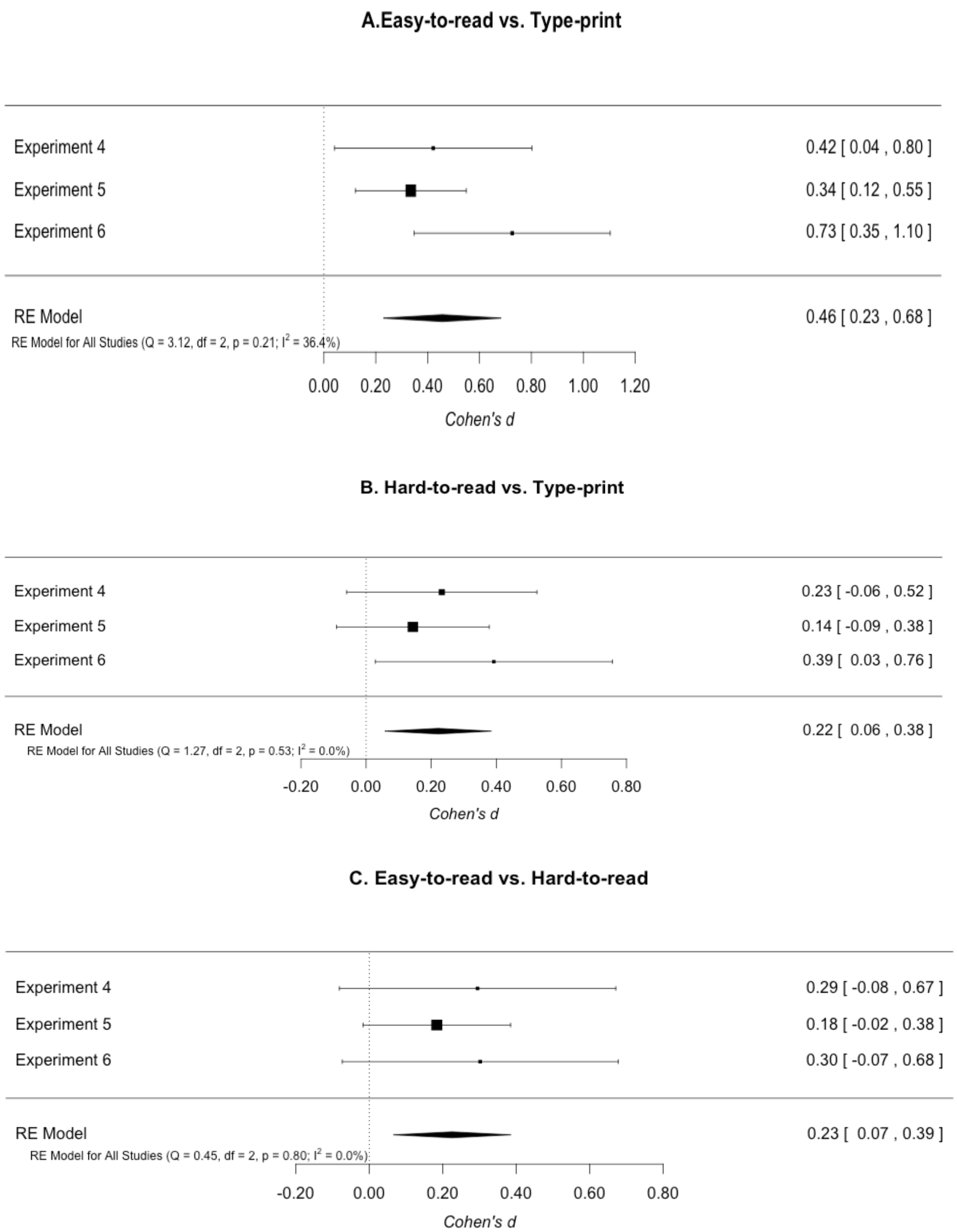

*Figure 10***.** Meta-analysis of Experiments 4, 5, and 6. A. Standardized mean difference of memory sensitivity between easy-to-read cursive and type-print. B. Standardized mean difference of memory sensitivity between hard-to-read cursive and type-print. C. Standardized mean difference of memory sensitivity between easy-to-read cursive and hard-to-read cursive.

CHAPTER 10. GENERAL DISCUSSION

One unexpected way to improve retention of material is to make the material more disfluent by changing the perceptual characteristics to make it hard-to-read. The mnemonic benefit afforded by perceptual disfluency is called the disfluency effect (Diemand-Yaueman, Oppenheimer, & Vaughan, 2011). Research on perceptual disfluency has yielded equivocal results, with some studies showing the effect (see Table 1) and others not (see Table 2). This has necessitated a search for potential boundary conditions of the disfluency effect (Oppenheimer & Atler, 2014).

The series of experiments reported in this dissertation examined the influence of perceptual disfluency on recognition memory using three disfluency manipulations: masking, blurring, and cursive.To better understand the theoretical underpinnings of the disfluency effect, I wanted to examine under what conditions disfluency does and does not constitute a desirable difficulty (Bjork, 1994). To this end, I examined several potential boundary conditions: Type of JOLs (i.e., item-by-item and aggregate), type of disfluency manipulation, encoding duration, list design, and the retention interval. In the following sections, I elaborate on the main findings of this dissertation.

Boundary Conditions

This dissertation identified several variables that potentially contribute to the heterogeneous findings regarding the disfluency effect. Experiment 1 tested whether type of JOLs (i.e., aggregate and item-by-item) and testing expectancy serve as potential moderating factors of the disfluency effect using masking, a commonly used perceptual disfluency manipulation. Three different groups were used. One group received instructions that alluded to a recognition memory test and were required to provide item-by-item JOLs during encoding. The second group received instructions that alluded to a recognition memory test and were required to provide aggregate JOLs after encoding the stimuli. The third group was not told about a memory test and was required to provide aggregate JOLs. Inclusion of three groups allowed for an examination of the independent influences of JOLs and testing expectancy on the disfluency effect.

Type of JOLs did not appear to moderate the disfluency effect; however, testing expectancy did moderate the disfluency effect. That is, a disfluency effect occurred when participants were not informed about an upcoming recognition memory test. This suggests that additional post-lexical processing that might be the result of preparing for a test can eliminate the disfluency effect. When participants are told about an immediate, upcoming memory test, all items receive additional processing that produce semantic representations that countervail the effects of perceptual disfluency. Indeed, under intentional conditions, the overall memory benefit was stronger. Because of this, no disfluency effect was found. In contrast, under incidental instructions, representations reflect the processing of the stimulus free of any intentional attempt to enhance the representations. Consistent with this, a disfluency effect was observed only when aggregate JOLs were used in conjunction with incidental instructions. In sum, the results of Experiment 1 suggest that a disfluency effect can be produced (with masking), but may be difficult to observe when there is high test expectancy. As a result of this, all subsequent experiments in this dissertation used incidental instructions with aggregate JOLs.

Another potentially important moderating factor examined was type of disfluency manipulation. In Experiment 2, a low-level blurring manipulation that putatively influences a pre-lexical locus was examined. In this experiment, studying blurred words for 500 ms did not engender superior recognition memory compared to clear words. It is important to note that while blurred words took longer to name, suggesting they are disfluent, accuracy was still very high (>98%). Thus, blurring does not increase difficulty that much. This is consistent with Yue et al.'s (2013) findings (also see Rosner et al., 2015, Experiment 3A).

The remaining experiments used a stronger perceptual disfluency manipulation—cursive. In Experiments 3, 4, 5, and 6, two levels of cursive were chosen: easy-to-read cursive and hard-to-read cursive. I chose two types of cursive because while both slow down processing and affect lexical processing, they do so to different degrees; hard-to-read cursive words take longer to recognize and are more error prone than easy-to-read cursive words, suggesting that hard-to-read cursive words exert a stronger influence on lexical processing (Perea et al., 2016a).

To examine the role of cursive, Experiment 3 had participants name easy-to-read cursive and hard-to-read cursive words, with a 500 ms encoding duration. Presenting the words in cursive did not produce a disfluency effect. In fact, hard-to-read cursive showed slightly poorer memory. To examine if the lack of disfluency effect was due to encoding duration to the stimuli themselves, I increased the encoding duration to 2 s in Experiment 4. With the longer encoding duration, a disfluency effect was found for both easy-to-read cursive words and hard-to-read cursive words. This cursive disfluency effect was replicated in Experiments 5 and 6. Specifically, in Experiment 5, a cursive disfluency effect was observed in both mixed and blocked lists. In Experiment 6, a similar pattern was observed after a 24-hour delay between study and test. Corroborating this pattern of results, a meta-analysis across Experiments 4, 5, and 6 showed a significant disfluency effect for easy-to-read (d = .46) and hard-to-read cursive (d = .22), with better memory for easy-to-read cursive words (d = .23).

Overall, not all perceptually disfluent stimuli are created equal. The discovery that both easy-to-read and hard-to-read cursive can produce better memory, but that the effect for easy-to-read cursive is larger in magnitude, suggests that *how* disfluency is manipulated matters because memory varies as a function of type of manipulation. It is possible that because hard-to-read cursive words are so difficult to recognize, which leads to increased lexical processing, that there is less post-lexical processing available to those stimuli. This will be discussed further in the section on mechanisms.

In sum, the experiments in this dissertation have identified three important moderators of the disfluency effect: testing expectancy, type of perceptual disfluency manipulation, and encoding duration. These results provide insights into the mechanisms behind the disfluency effect, which are discussed next.

Mechanisms of the Disfluency Effect

Chapter 1 identified three loci in the word recognition processing that may be responsible for the disfluency effect: pre-lexical, lexical, and post-lexical. Of interest is how the findings of this dissertation fit with accounts that place the disfluency effect at each of these loci.

To recapitulate, a pre-lexical account would posit that disfluency enhances memory because of noisy mapping of features to letters that occurs early on in processing. Lexical accounts, such as the compensatory processing account (Mulligan, 1996), posit that disfluency enhances memory because of increased interaction between the letter-level and word-level during word recognition. Lexical accounts would predict that manipulations that result in increased interactivity between the letter-level and word-level should elicit better memory. Lastly, post-lexical accounts (e.g., Atler et al., 2007) posit that the subjective experience of disfluency fosters control processes that enhance memory by way of increased conceptual or semantic processing.

Examining whether the disfluency effect arises at a pre-lexical locus, Experiment 2 employed a manipulation that affects a pre-lexical locus: low-level blurring (Reingold & Rayner, 2007). No disfluency effect occurred when blurring was used with a sufficient encoding duration that allowed for word recognition (accuracy was at ceiling), thereby providing evidence that the disfluency effect does not arise solely at a pre-lexical locus.

To examine whether the disfluency effect arises at a lexical locus, cursive words were used as they putatively influence this locus during encoding (Barnhart & Goldinger, 2010; Perea et al., 2016a, 20016b). A lexical account (i.e., compensatory processing; Mulligan, 1996) would predict cursive words to produce a disfluency effect, which was what was observed in Experiments 4, 5, and 6. Examining easy-to-read and hard-to-read cursive, a lexical account would predict a larger memory benefit for hard-to-read cursive compared to easy-to-read cursive due to the fact that hard-to-read cursive engenders greater top-down processing than easy-to-read cursive. Contrary to this, in Experiment 3, recognition memory for hard-to-read cursive tended to be worse than easy-to-read cursive. The meta-analysis confirmed this: While an overall disfluency effect was observed with hard-to-read cursive (d = .22), the effect was smaller than the disfluency effect observed with easy-to-read cursive words (d = .46). Further, easy-to-read cursive words were better remembered than hard-to-read cursive words. Taken together, this suggests that locus of the disfluency effect is not strictly lexical in nature.

Examining a post-lexical locus, Experiment 1 showed that a manipulation that is subserved by a post-lexical locus (i.e., high testing expectancy) attenuated the disfluency effect.

Furthermore, in Experiment 4, increasing the encoding duration from 500 ms to 2 s, which allows for more lexical and post-lexical processing, resulted in a disfluency effect for both easy-to-read and hard-to-read cursive stimuli. This provides some evidence that the disfluency effect is affected by post-lexical factors.

Although the aforementioned results suggest that the disfluency effect is post-lexical, this is not the whole story. Post-lexical accounts that posit that perceptual disfluency acts solely as a metacognitive (subjective) cue also would posit that there is only a quantitative difference between perceptual disfluency manipulations. That is, a memory benefit should be observed independent of the perceptual disfluency manipulation used. Thus, a disfluency effect should be present for both easy-to-read cursive and hard-to-read cursive words. While we did find an overall cursive effect for both easy-to-read and hard-to-read cursive in Experiments 4, 5, and 6, the meta-analysis revealed that the disfluency effect was larger for easy-to-read cursive words.

These results suggest that the disfluency effect occurs when perceptual disfluency affects processing during word recognition thereby triggering increased post-lexical processing. However, because post-lexical resources (e.g., working memory and attention control) are limited (Evans & Stanovich, 2013), if the word is too hard to recognize (due to increased lexical processing), the benefits may be outweighed by an increase in cognitive load (Sweller, 1994), producing smaller or weaker memory effects. This is what was observed with hard-to-read cursive. Corroborating this, Weissgerber and Reinhard (2017) examined the disfluency effect using a short passage on language lateralization in the brain. In this passage, they used two disfluency manipulations: Half the participants saw the passage in an atypical font and the other half saw the passage where 20% of the words had their letters scrambled or transposed (i.e., nonwords such as *jugde* had to be interpreted as *judge*). The scrambled letter manipulation was assumed to affect lexical-level processing. While both disfluency manipulations produced a disfluency effect, the atypical font benefited recall more than the lexical manipulation. Weissgerber and Reinhard reasoned that descrambling the words placed greater demands on working memory capacity, which left fewer post-lexical resources available while processing the passage. Thus it seems that perceptual disfluency is most beneficial to memory when the perceptual disfluency manipulation does not expend too many post-lexical resources.

The results from this dissertation implicate not just one locus, but three loci underlying the disfluency effect (see Figure 11). All perceptual disfluency manipulations influence a pre-lexical locus. That is, in order to recognize a perceptual disfluent word, the pre-lexical locus is activated. With handwritten cursive, there appears to be an initial encoding cost (i.e., increased naming latency and lower accuracy rates) that arises at a pre-lexical locus when mapping letter features to abstract letter representations. In addition, both easy-to-read and hard-to-read cursive words yield an additional cost at a lexical locus, with hard-to-read cursive more strongly affecting the rate at which evidence is accumulated to know what the word is (Perea et al., 2016a, 2016b). This is bolstered by the finding that hard-to-read cursive words showing longer naming latencies and lower accuracy rates than easy-to-read cursive. The perceptual disfluency that is experienced causes increased post-lexical processing. The amount of post-lexical processing a word receives, however, is modulated by how difficult the stimulus is to recognize. If the stimulus is too hard to process, more resources are needed to recognize the word at the lexical level (Geller, Still, & Morris, 2015), which leaves less post-lexical resources. This leads to a weaker or smaller memory benefits.

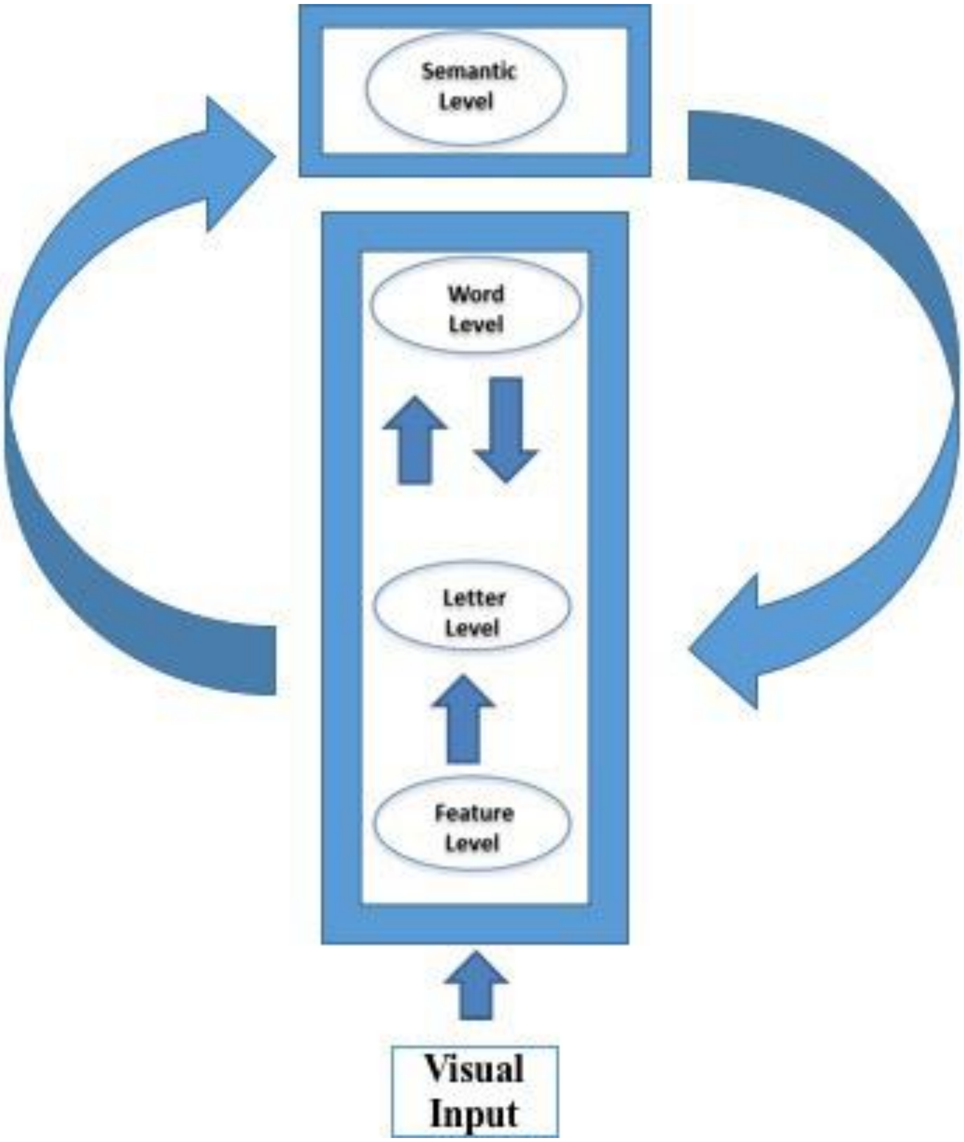

Figure 11. This diagram sketches out the mechanisms underlying the disfluency effect. The disfluency effect appears to be the result of interactivity between word recognition and semantic processing. In order to get the disfluency effect, the perceptual manipulation must affect the rate at which lexical evidence accumulates. As a result of this perceptual disfluency, this stimulus gets deeper processing. However, if the manipulation strongly affects the lexical level, there will be fewer post-lexical resources allocated to processing the word leading to a weaker, less durable, memory representation.

Related to the aforementioned, the enhancement of memory due to perceptual disfluency has recently been contextualized within the conflict monitoring framework of cognitive control (Rosner et al., 2015). In the conflict monitoring framework (Botvinick, Braver, Barch, Carter, & Cohen, 2001), monitoring and recruitment of attentional control is modulated by the amount of conflict experienced during processing. Attentional control is carried out in the brain via activation of the anterior cingulate cortex (ACC), which detects conflict during processing. The activation of the ACC engenders increased activation of the dorsolateral prefrontal cortex (DLPF)—an area important for sustained or selective attention. The conflict monitoring model has been primarily used within the context of response inhibition (e.g., Stroop task) and the detection of errors. It is possible that during the course of word processing, perceptual disfluency increases response ambiguity, which in turn activates the ACC and increases cognitive control mechanisms leading to a memory benefit. This fits well with the current set of results implicating an interaction between word recognition and post-lexical processing.

Metamemory Judgments and Disfluency

Although my primary interest was actual memory performance, I also assessed participants' metamemory through JOLs. One reason I collected JOLs was to see if they moderated the disfluency effect. A second reason why I collected JOLs was because individuals are presumed to judge their own learning and JOLs play a large role in controlling and regulating future study behavior (e.g., Son & Metcalfe, 2000). Thus, it is important to examine how perceptual disfluency influences JOLs.

The consensus is that students have poor metamemory skills (Bjork et al., 2013). Usually students are unable to accurately monitor and regulate their own learning and thus show misalignment or miscalibration between what they think they know and what they actually know. This is problematic as students may not adopt and implement effective learning strategies to improve learning (Hartwig & Dunlosky, 2012; McCabe, 2012). One cue that contributes to this misalignment is fluency.

Post-lexical metacognitive theories make the assumption that making something disfluent should lead to better monitoring, which in turn, should produce better control and regulation in learning (Pierger, Mengelkamp, & Bannert, 2016). That is, individuals use perceived disfluency as a metacognitive cue to evoke better control and regulatory processes. Contrary to this, various perceptual disfluency manipulations such as blurring (Yue et al., 2013) and small font size (Rhodes & Castel, 2008) do not produce changes in actual memory performance despite differences in metamemory judgments (i.e., lower JOLs). This discrepancy between predicted and actual memory performance is taken as evidence that individuals fail to engage in adequate regulatory and monitoring processes during learning. Is this in accord with the results obtained here?

In almost all reported experiments, individuals predicted that they would recognize more fluent stimuli than disfluent stimuli: nonmasked vs. masked (Experiment 1), clear vs. blurred (Experiment 2), and type-print vs. cursive (Experiments 3-6). This pattern of results is in accordance with the fluency hypothesis (Rhodes & Castel, 2008, 2009) and also the *ease of processing heuristic* (Kornell, Rhodes, Castel, & Tauber, 2011), which posit that individuals judge stimuli as better remembered when information seems easy to process. Participants in the dissertation experiments predicted that they would have better memory for fluent stimuli; however, in the majority of studies (i.e., Experiments 1, 3, 4, 5, and 6), metamemory judgments were lower for disfluent stimuli, but actual memory was better. Thus, individuals are influenced by disfluency, which leads to better control and regulatory processes during study, producing better memory.

Participants' metamemory (aggregate JOLs) accuracy was generally lower for disfluent words compared to fluent words. This presumably lead to better regulatory and monitoring processes producing the disfluency effect (Pierger et al., 2016). However, there is some evidence to suggest that JOLs do not reflect experiential processing (i.e., how long it actually takes them to name words), but instead reflect belief-based processing (i.e., things are learned better when they are easier to read) (Mueller et al., 2014). If so, JOLs may not have caused better performance in the current studies.

Post-lexical or metacognitive accounts of disfluency (e.g., Alter, 2013) would predict that hard-to-read cursive and easy-to-read cursive should produce similar metamemory judgments as both are disfluent—that is, both engender longer latencies and lower accuracy. This prediction was not supported by the current data. In the case of cursive, in several experiments (4,5, and 6) participants judged hard-to-read cursive words as being more disfluent than easy-to-read cursive. From a metacognitive perspective, then, participants should have assigned more post-lexical resources to hard-to-read cursive words than easy-to-read cursive words, which would have produced better memory. On an actual memory test, however, easy-to-read cursive words were better remembered than words printed in hard-to-read cursive. Thus, rather than metamemory judgements guiding performance, it is possible that processing disfluency mediates better recognition memory performance. To support this, Sanchez and Jager (2012) showed that reading times (on-line measure of difficulty) of a passage presented in an atypical font mediated performance whereas ratings of difficulty (i.e., perceived difficulty) do not. Thus, it appears that a simple metacognitive theory of disfluency is not tenable to explain the current pattern of results. Nevertheless, the pattern of results aligns well with the cognitive monitoring framework (Botvinick et al., 2001) and the depth of processing mechanism outlined here. That is, the memory benefit observed for disfluent words is a byproduct of processing disfluency. Participants may have allocated more attention to hard-to-read cursive words, but given how objectively difficult hard-to-read cursive words are and the encoding duration, it is possible that participants could not process them to the level needed to produce better memory. Future research should examine the relationship between perceptual disfluency and metamemory judgments is needed.

Practical Implications

Recent studies by Diemand-Yauman et al. (2010) and French et al. (2013) have recommended that teachers and students use disfluency to enhance learning. Is this recommendation premature? Although the potential for disfluent stimuli to improve learning is great, evidence is mixed. Using results from laboratory studies, one can begin to examine whether perceptual disfluency might have practical implications.

One common criticism raised against disfluency research is that the manipulations used in laboratory based experiments are not realistic (Serra, 2016). How often are students confronted with passages that are out of focus or inverted? Further, when teachers present material in-class, how often do they use hard-to-read fonts? The answer to both of these questions is not often. Across several experiments, I used a manipulation that is both naturally disfluent and educationally realistic: handwritten cursive. Teachers often use handwritten cursive when creating notes, writing on transparencies or whiteboards, or giving students written feedback on their work. Further, students are likely to study from handwritten notes (their own or those of another student). Across several experiments presenting words in easy-to-read and hard-to-read cursive enhanced memory compared to type-print, with the effect size being small for hard-to-read cursive stimuli and moderate for easy-to-read cursive. This demonstrates that in the laboratory, the disfluency effect can be obtained with a more realistic manipulation. Further research should examine whether cursive can provide the same memory benefit in a classroom setting.

In addition, I examined two conditions that serve to approximate real-life learning and studying situations. Experiment 5 examined if list format (blocked vs. mixed) matters when studying disfluent words. Students are rarely exposed to both fluent and disfluent information intermixed at time of study. If the disfluency effect only occurs when distinctive or unusual stimuli are studied against a backdrop of common or typical stimuli, this would limit the practical implications. What this would mean is that the benefit of disfluency only arises very early on in learning when the material is distinctive and attracts attention. As learning progresses, students would get used to the manipulation, leading to the disappearance of any memory benefit. Results revealed that practice format had no effect on the memory benefit conferred. The magnitude of the effect was statistically comparable across both mixed and blocked lists. This finding suggests that it does not matter how one studies disfluent (cursive) words. Studying disfluent words aids memory, and this enhancement can be detected when incidental instructions are used.

Finally, in a real classroom environment, students are rarely given tests immediately after studying material. Tests may be given hours, days, or weeks after exposure to the material. To emulate this set of circumstances in the laboratory, I gave participants a recognition memory test 24-hours after studying cursive and type-print words. Even after a 24-hour delay, the effects of cursive persisted. It is important to point out that while students were not explicitly told about an upcoming memory test, students might have adopted an intentional strategy. That is, they could have figured out there was going to be a test when they were asked to participate in a study that had two parts that were 24 hours apart. If this is true, intentional instructions may not moderate the disfluency effect after a delay. Future research should examine this with intentional instructions, or ask students about their expectations at the end of the first part. Nonetheless, it has been shown that disfluency can benefit memory in short-term and long-term contexts.

The results show that cursive can be used to increase short- and long-term learning, and that this effect does not depend on the practice format. However, there is some evidence that limits the utility of disfluency. In Experiment 1, I demonstrated that the disfluency effect is modulated by testing expectancy. That is, when individuals were told about an upcoming memory test, disfluency did not enhance memory. In an education setting, students are usually told about a memory test. Thus, disfluency might not be an effective manipulation to enhance memory. However, it is important to note that the effect of instructions was examined with masking, and not cursive. It is not clear if cursive would also show this same pattern of effects. As stated in the discussion of Experiment 5, masking may rely on different mechanisms than cursive. For instance, the effect of masking is weakened in a blocked design (Westerman & Greene, 1997) whereas I found no difference between blocked and mixed designs for cursive words in this dissertation. Thus, a disfluency effect driven by distinctiveness may be countervailed by testing expectancy. In contrast, a disfluency effect driven by depth of processing may not be affected as much by intent to learn. Future studies should examine whether testing expectancy influences the disfluency effect observed with cursive.

The present studies focused on the desirable effect of disfluency using simple word lists. In a classroom environment, almost all learned material is complex. Future studies should examine the effect of cursive on comprehension and transfer with more complex stimuli.

Finally, before implementation in a classroom setting, the disfluency effect should be examined with other types of memory tests. The disfluency effect has largely been found using a recognition test (e.g., Hirshman & Mulligan, 1991; Mulligan, 1996, 1999) rather than a free recall test (e.g., Hirshman & Mulligan, 1991; Narine, 1988). Indeed, finding a disfluency effect in free recall has been mixed (Narine, 1988; Strukelj et al., 2015; Yue et al., 2012). This could be due to disfluency benefiting tests that are sensitive to the bottom-up and top-down processes involved with forming an item-specific representation (i.e., recognition). Moreover, disfluency might not be beneficial on tests that rely on retrieval pathways (i.e., free recall). Future studies should examine whether the disfluency effect can be found in free recall. That is, it need to be determined whether the effects of disfluency are generalizable across different tasks.

Conclusion

This dissertation demonstrated that perceptual disfluency can produce a desirable difficulty effect, with better immediate (Experiments 1, 4, and 5) and delayed (Experiment 6) recognition memory performance for perceptually disfluent items. However, there are important moderators of the disfluency effect. First, the inclusion of intentional instructions appears to eliminate the disfluency effect on an immediate test (Experiment 1). Second, the disfluency effect requires words to be studied for a sufficient duration, allowing for increased post-lexical processing (Experiment 4). Third, the type of disfluency manipulation used matters (Experiments 4-6).

Current thinking on the disfluency effect suggests the locus is solely post-lexical in nature (Alter, 2013; Kuhl & Eitel, 2016). From this perceptive, the perception of disfluency (regardless of type of disfluency manipulation) evokes increased post-lexical processing that produces the disfluency effect. The current set of results suggests otherwise. Using three different perceptual manipulations (i.e., masking, blurring, and cursive), I have shown that the disfluency effect can be modulated by the type of perceptual disfluency manipulation used. A perceptually disfluent manipulation that taps a pre-lexical locus (blurring) did not produce a disfluency effect. However, manipulations that tap a higher, lexical, locus (i.e., masking and cursive) did produce a disfluency effect. This strongly suggests that perceptual disfluency manipulations that affect the later stages of word recognition are needed to produce the disfluency effect. Furthermore, perceptual disfluency manipulations that tap the lexical locus to different degrees modulate the memory benefit. The use of a strong perceptual disfluency manipulation like hard-to-read cursive requires a lot more lexical processing than easy-to-read cursive. Because of this, hard-to-read cursive words receive less post-lexical processing, thereby weakening the memory benefit. Current theories of disfluency need to be modified to account for these findings.

Researchers examining the disfluency effect should keep this in mind while planning future studies. It is recommended that researches not use intentional manipulations that countervail the effects of disfluency at time of study (e.g., intentional instructions or item-by-item JOLs). Further, researchers should ensure that disfluent stimuli are sufficiently processed. Finally, the perceptual manipulations used should be sufficiently disfluent. The perceptual manipulation must be sufficiently strong to make encoding difficult, but not so strong that it consumes too many attentional resources needed for post-lexical processing. To control for this, researchers could employ finer grained methods such as eye tracking, or mathematical modeling (e.g., distributional analyses) to examine the locus of a perceptual manipulation.

Overall, this study makes important contributions to research on disfluency and desirable difficulties. The findings help to illuminate inconsistent disfluency effects on a broader level because I have shown that they are not generalizable across testing expectancy, encoding duration, and type of disfluency manipulation. Whether or when disfluency is desirable depends upon multiple factors. The disfluency effect is not as straightforward as placing something in a hard-to-read font.

REFERENCES

Alban, M. W., & Kelley, C. M. (2013). Embodiment meets metamemory: Weight as a cue for metacognitive judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(5), 1628.

Alter, A. L. (2013). The benefits of cognitive disfluency. *Current Directions in Psychological Science*, *22*(6), 437–442. https://doi.org/10.1177/0963721413498894

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. https://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.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. *The Psychology of Learning and Motivation*, *2*, 89–195.

Baddeley, A. (1981). The concept of working memory: A view of its current state and probable future development. *Cognition*, *10*(1), 17–23.

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.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 575–589. https://doi.org/10.1016/S0022-5371(75)80045-4

Balota, D. A., & Spieler, D. H. (1999). Word frequency, repetition, and lexicality effects in word recognition tasks: Beyond measures of central tendency. *Journal of Experimental Psychology: General*, *128*(1), 32-55.

Balota, D. A., & Yap, M. J. (2011). Moving beyond the mean in studies of mental chronometry the power of response time distributional analyses. *Current Directions in Psychological Science, 20*(3), 160–166.

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., … Treiman, R. (2007). The English lexicon project. *Behavior Research Methods, 39*(3), 445–459.

Barnhart, A. S., & Goldinger, S. D. (2010). Interpreting chicken-scratch: lexical access for handwritten words. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(4), 906-923.

Barnhart, A. S., & Goldinger, S. D. (2013). Rotation reveals the importance of configural cues in handwritten word perception. *Psychonomic Bulletin & Review*, *20*(6), 1319–1326.

Barnhart, A. S., & Goldinger, S. D. (2015). Orthographic and phonological neighborhood effects in handwritten word perception. *Psychonomic Bulletin & Review*, *22*(6), 1739–1745.

Begg, I., Duft, S., Lalonde, P., Melnick, R., & Sanvito, J. (1989). Memory predictions are based on ease of processing. *Journal of Memory and Language*, *28*(5), 610–632.

Besken, M., & Mulligan, N. W. (2013). Easily perceived, easily remembered? Perceptual interference produces a double dissociation between metamemory and memory performance. *Memory & Cognition*, *41*(6), 897–903.

Besken, M., & Mulligan, N. W. (2014). Perceptual fluency, auditory generation, and metamemory: analyzing the perceptual fluency hypothesis in the auditory modality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(2), 429-440.

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., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, *64*, 417–444.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652.

Brandt, M. J., IJzerman, H., Dijksterhuis, A., Farach, F. J., Geller, J., Giner-Sorolla, R., … Van't Veer, A. (2014). The replication recipe: What makes for a convincing replication? *Journal of Experimental Social Psychology*, *50*, 217–224.

Braver, S. L., Thoemmes, F. J., & Rosenthal, R. (2014). Continuously cumulating meta-analysis and replicability. *Perspectives on Psychological Science*, *9*(3), 333–342. https://doi.org/10.1177/1745691614529796

Carpenter, S. K., Pashler, H., & Cepeda, N. J. (2009). Using tests to enhance 8th grade students' retention of U.S. history facts. *Applied Cognitive Psychology*, *23*(6), 760–771. https://doi.org/10.1002/acp.1507

Carpenter, S. K., Pashler, H., Wixted, J. T., & Vul, E. (2008). The effects of tests on learning and forgetting. *Memory & Cognition*, *36*(2), 438–448.

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

Coltheart, M., Davelaar, E., Jonasson, J. F., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention & Performance VI* (pp. 535-555). Hillsdale, NJ: Erlbaum

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684.

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, *25*(1), 7–29. https://doi.org/10.1177/0956797613504966

Davis, C. J. (2003). Factors underlying masked priming effects in competitive network models of visual word recognition. In S. Kinoshita & S. J. Lupker (Eds.), *Masked Priming: The State of the Art* (pp. 121–170), Hove, England: Psychology Press.

De Zuniga, C. M., Humphreys, G. W., & Evett, L. J. (1991). Additive and interactive effects of repetition, degradation, and word frequency in the reading of handwriting. *Basic Processes in Reading: Visual Word Recognition*, 10–33.

Diemand-Yauman, C., Oppenheimer, D. M., & Vaughan, E. B. (2011). Fortune favors the (): Effects of disfluency on educational outcomes. *Cognition*, *118*(1), 111–115.

Dunlosky, J., & Nelson, T. O. (1992). Importance of the kind of cue for judgments of learning (JOL) and the delayed-JOL effect. *Memory & Cognition*, *20*(4), 374–380.

Eitel, A., & Kühl, T. (2016). Effects of disfluency and test expectancy on learning with text. *Metacognition and Learning*, *11*(1), 107–121. https://doi.org/10.1007/s11409-015-9145-3

Eitel, A., Kühl, T., Scheiter, K., & Gerjets, P. (2014). Disfluency meets cognitive load in multimedia learning: Does harder-to-read mean better-to-understand? *Applied Cognitive Psychology*, *28*(4), 488–501.

Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: advancing the debate. *Perspectives on Psychological Science*, *8*(3), 223–241. https://doi.org/10.1177/1745691612460685

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.

Fiedler, K. (2013). Fluency and behavior regulation: Adaptive and maladaptive consequences of a good feeling. In C. Unkelback & R. Greifeneder (Eds.), *The experience of thinking: How the fluency of mental processes influences cognition and behavior* (pp. 234-255) Hove, England: Psychology Press.

Foster, N. L., & Sahakyan, L. (2012). Metacognition influences item-method directed forgetting. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *38*(5), 1309–1324.

French, M. M. J., Blood, A., Bright, N. D., Futak, D., Grohmann, M. J., Hasthorpe, A., … Tabor, J. (2013). Changing fonts in education: how the benefits vary with ability and dyslexia. *The Journal of Educational Research*, *106*(4), 301–304.

Geller, J., Still, M. L., & Morris, A. L. (2016). Eyes wide open: Pupil size as a proxy for inhibition in the masked-priming paradigm. *Memory & Cognition*, *44*(4), 554–564. https://doi.org/10.3758/s13421-015-0577-4

Glanzer, M., & Adams, J. K. (1990). The mirror effect in recognition memory: data and theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(1), 5-16.

Glenberg, A. M. (1976). Monotonic and nonmonotonic lag effects in paired-associate and recognition memory paradigms. *Journal of Verbal Learning and Verbal Behavior*, *15*(1), 1–16. https://doi.org/10.1016/S0022-5371(76)90002-5

Gomez, P., & Perea, M. (2014). Decomposing encoding and decisional components in visual-word recognition: A diffusion model analysis. *The Quarterly Journal of Experimental Psychology*, *67*(12), 2455–2466.

Graf, P., & Levy, B. A. (1984). Reading and remembering: Conceptual and perceptual processing involved in reading rotated passages. *Journal of Verbal Learning and Verbal Behavior*, *23*(3), 405–424. http://doi.org/10.1016/S0022-5371(84)90281-0

Hartwig, M. K., & Dunlosky, J. (2012). Study strategies of college students: Are self-testing and scheduling related to achievement? *Psychonomic Bulletin & Review*, *19*(1), 126-134.

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. http://doi.org/10.1037/0278-7393.17.3.507

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.

Hu, X., Li, T., Zheng, J., Su, N., Liu, Z., & Luo, L. (2015). How much do metamemory beliefs contribute to the font-size effect in judgments of learning? *PLoS ONE*, *10*(11). http://doi.org/10.1371/journal.pone.0142351

Jia, X., Li, P., Li, X., Zhang, Y., Cao, W., Cao, L., & Li, W. (2016). The effect of word frequency on judgments of learning: Contributions of beliefs and processing fluency. *Frontiers in Psychology*, *6*. http://doi.org/10.3389/fpsyg.2015.01995

Kahneman, D. (2011). *Thinking, fast and slow* (First edition.). New York: Farrar, Straus and Giroux.

Karpicke, J. D., Butler, A. C., & Roediger, H. L. III. (2009). Metacognitive strategies in student learning: Do students practice retrieval when they study on their own? *Memory*, *17*(4), 471-479

Kinoshita, S. (1989). Generation enhances semantic processing? The role of distinctiveness in the generation effect. *Memory & Cognition*, *17*(5), 563–571

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, *126*(4), 349-370.

Koriat, A., & Ma'ayan, H. (2005). The effects of encoding fluency and retrieval fluency on judgments of learning. *Journal of Memory and Language*, *52*(4), 478–492. http://doi.org/10.1016/j.jml.2005.01.001

Kühl, T., & Eitel, A. (2016). Effects of disfluency on cognitive and metacognitive processes and outcomes. *Metacognition and Learning*, *11*(1), 1–13.

Larsen, D. P., Butler, A. C., & Roediger, H. L. (2009). Repeated testing improves long-term retention relative to repeated study: a randomised controlled trial. *Medical Education*, *43*(12), 1174–1181. https://doi.org/10.1111/j.1365-2923.2009.03518.x

Lakens, D., & Etz, A. J. (in press). Too true to be bad: When sets of studies with significant and nonsignificant findings are probably true. *Social Psychological and Personality Science*, 1948550617693058. https://doi.org/10.1177/1948550617693058

Lupker, S. J., & Davis, C. J. (2009). Sandwich priming: A method for overcoming the limitations of masked priming by reducing lexical competitor effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(3), 618.

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. Psychology press.

Magreehan, D. A., Serra, M. J., Schwartz, N. H., & Narciss, S. (2016). Further boundary conditions for the effects of perceptual disfluency on judgments of learning. Metacognition and Learning, *11*(1), 35–56. https://doi.org/10.1007/s11409-015-9147-1

Manso De Zuniga C., Humphreys G.W., & Evett L.J. (1991) Additive and interactive effects of repetition, degradation, and word frequency in the reading of handwriting. In D Besner and G.W. Humphreys (eds). *Basic processes in reading: Visual word recognition* (10-33). Hillsdale, NJ: Erlbaum.

Matvey, G., Dunlosky, J., & Guttentag, R. (2001). Fluency of retrieval at study affects judgments of learning (JOLs): An analytic or nonanalytic basis for JOLs? *Memory & Cognition*, 29(2), 222–233. https://doi.org/10.3758/BF03194916

McShane, B. B., & Böckenholt, U. (2017). Single-paper meta-analysis: Benefits for study summary, theory testing, and replicability. *Journal of Consumer Research*, 43(6), 1048–1063. https://doi.org/10.1093/jcr/ucw085

Matzke, D., & Wagenmakers, E.-J. (2009). Psychological interpretation of the ex-Gaussian and shifted Wald parameters: A diffusion model analysis. *Psychonomic Bulletin & Review*, *16*(5), 798–817.

McCabe, J. (2011). Metacognitive awareness of learning strategies in undergraduates. *Memory & Cognition*, *39*(3), 462-476.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, *88*(5), 375.

Metcalfe, J., & Finn, B. (2008). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin & Review*, *15*(1), 174–179

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Reason*, *4*(2), 61–64.

Mueller, M. L., Dunlosky, J., Tauber, S. K., & Rhodes, M. G. (2014). The font-size effect on judgments of learning: Does it exemplify fluency effects or reflect people's beliefs about memory? *Journal of Memory and Language*, *70*, 1–12.

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. http://doi.org/10.1037/0278-7393.22.5.1067

Mulligan, N. W. (1998). Perceptual interference at encoding enhances recall for high-but not low-imageability words. *Psychonomic Bulletin & Review*, *5*(3), 464–469.

Mulligan, N. W., Susser, J. A., & Smith, S. A. (2016). The testing effect is moderated by experimental design. *Journal of Memory and Language*, *90*, 49–65. https://doi.org/10.1016/j.jml.2016.03.006

Nairne, J. S. (1988). The mnemonic value of perceptual identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(2), 248.

Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: measures of effect size for some common research designs. *Psychological Methods*, *8*(4), 434–447.

Oppenheimer, D. M. (2008). The secret life of fluency. *Trends in Cognitive Sciences*, *12*(6), 237–241.

Oppenheimer, D. M., & Alter, A. L. (2014). The search for moderators in disfluency research. *Applied Cognitive Psychology*, *28*(4), 502–504. http://doi.org/10.1002/acp.3023

Ozubko, J. D., & Macleod, C. M. (2010). The production effect in memory: Evidence that distinctiveness underlies the benefit. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *36*(6), 1543–1547. https://doi.org/10.1037/a0020604

Perea, M., Gil-López, C., Beléndez, V., & Carreiras, M. (2016). Do handwritten words magnify lexical effects in visual word recognition? *Quarterly Journal of Experimental Psychology* (2006), *69*(8), 1631–1647. https://doi.org/10.1080/17470218.2015.1091016

Perea, M., Marcet, A., Uixera, B., & Vergara-Martínez, M. (2016). Eye movements when reading sentences with handwritten words. *Quarterly Journal of Experimental Psychology* (2006), 1–9. https://doi.org/10.1080/17470218.2016.1237531

Pieger, E., Mengelkamp, C., & Bannert, M. (2016). Metacognitive judgments and disfluency – Does disfluency lead to more accurate judgments, better control, and better performance? *Learning and Instruction*, *44*, 31–40. https://doi.org/10.1016/j.learninstruc.2016.01.012

Qiao, E., Vinckier, F., Szwed, M., Naccache, L., Valabrègue, R., Dehaene, S., & Cohen, L. (2010). Unconsciously deciphering handwriting: subliminal invariance for handwritten words in the visual word form area. *Neuroimage*, *49*(2), 1786–1799.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, *105*(1), 125–157.

Reingold, E. M., & Rayner, K. (2006). Examining the word identification stages hypothesized by the EZ Reader model. *Psychological Science*, *17*(9), 742–746.

Rhodes, M. G., & Castel, A. D. (2008). Memory predictions are influenced by perceptual information: evidence for metacognitive illusions. *Journal of Experimental Psychology: General*, *137*(4), 615-625.

Rhodes, M. G., & Castel, A. D. (2009). Metacognitive illusions for auditory information: Effects on monitoring and control. *Psychonomic Bulletin & Review*, *16*(3), 550–554.

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*(3), 249–255. https://doi.org/10.1111/j.1467-9280.2006.01693.x

Rosner, T. M., Davis, H., & Milliken, B. (2015). Perceptual blurring and recognition memory: A desirable difficulty effect revealed. *Acta Psychologica*, *160*, 11–22.

Rummer, R., Schweppe, J., & Schwede, A. (2015). Fortune is fickle: null-effects of disfluency on learning outcomes. *Metacognition and Learning*, 1–14.

Sanchez, C. A., & Jaeger, A. J. (2015). If it's hard to read, it changes how long you do it: Reading time as an explanation for perceptual fluency effects on judgment. *Psychonomic Bulletin & Review*, *22*(1), 206–211.

Schneider, W., Eschmann, A., & Zuccolotto, A. (2002). E-Prime v1. 1. *Pittsburgh, PA: Psychology Software Tools Inc*.

Serra, M. (2016, April 14). Does Processing Fluency Really Matter for Metacognition in Actual Learning Situations? (Part Two). Retrieved June 13, 2017, from http://www.improvewithmetacognition.com/processing-fluency-really-matter-metacognition-actual-learning-situations/

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

Strukelj, A., Scheiter, K., Nyström, M., & Holmqvist, K. (2015). Exploring the lack of a disfluency effect: Evidence from eye movements. *Metacognition and Learning*, 1–18.

Sungkhasettee, V. W., Friedman, M. C., & Castel, A. D. (2011). Memory and metamemory for inverted words: illusions of competency and desirable difficulties. *Psychonomic Bulletin & Review*, *18*(5), 973–978. http://doi.org/10.3758/s13423-011-0114-9

Susser, J. A., Jin, A., & Mulligan, N. W. (2016). Identity priming consistently affects perceptual fluency but only affects metamemory when primes are obvious. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *42*(4), 657–662.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. Learning and Instruction, *4*(4), 295–312. https://doi.org/10.1016/0959-4752(94)90003-5

Weissgerber, S. C., & Reinhard, M.-A. (2017). Is disfluency desirable for learning? *Learning and Instruction*, *49*, 199–217. https://doi.org/10.1016/j.learninstruc.2017.02.004

Weltman, D., & Eakin, M. (2014). Incorporating unusual fonts and planned mistakes in study materials to increase business student focus and retention. *INFORMS Transactions on Education*, *15*(1), 156–165.

Westerman, D. L., & Greene, R. L. (1997). The effects of visual masking on recognition: Similarities to the generation effect. *Journal of Memory and Language*, *37*(4), 584–596.

Wong, Y. K., Twedt, E., Sheinberg, D., & Gauthier, I. (2010). Does Thompson’s Thatcher Effect reflect a face-specific mechanism? *Perception*, *39*(8), 1125–1141.

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.

APPENDIX A: STUDY MATERIALS

The 200 words used in each experiment.

alto	axis	atom	ball
army	bath	babe	beer
auto	bell	bean	bird
bass	boat	belt	bout
bees	bulb	body	bush
cafe	cage	bull	city
cell	cent	cape	work
clan	clay	chow	corn
cone	kiwi	coal	dean
cult	deer	cord	desk
debt	knot	odor	disc
dice	door	dirt	dude
disk	duty	fate	ears
duke	fame	flaw	fees
exam	fist	font	flea
feet	folk	frog	food
flux	fort	girl	fury
fork	gift	grid	goal
gene	data	hair	gulf
golf	heck	heel	hall
guru	herb	hero	helm
life	hive	hood	hill
hemp	hour	icon	horn
hose	item	jazz	info
iris	hoop	jury	aunt
king	lady	lake	kind
lamp	lane	luck	lamb
liar	lion	mall	lens
loss	lots	maze	lore
lust	lynx	menu	lung
jeep	math	mode	meal
meat	memo	myth	mice
mile	mill	none	monk
mood	moss	path	noon
newt	node	pond	oath

norm	nova	reed	poem
poet	oven	role	punk
dune	pole	scum	reef
rice	rand	soda	roof
kite	road	spec	self
time	sale	tale	song
sync	silk	text	stud
tech	soul	town	tear
tome	tent	unit	tomb
tree	tool	vise	tray
vice	tube	watt	vale
wand	visa	wood	volt
wife	wine	yarn	week
puck	yard	area	word
ammo	amps	aura	zone

APPENDIX B: IRB APPROVAL

IOWA STATE UNIVERSITY
OF SCIENCE AND TECHNOLOGY

Institutional Review Board
Office for Responsible Research
Vice President for Research
2420 Lincoln Way, Suite 202
Ames, Iowa 50014
515 294-4566

Date: 9/7/2016

To: Jason Geller
W112 Lagomarcino

CC: Dr. Veronica Dark
W112 Lagomarcino Hall

From: Office for Responsible Research

Title: Name that Word

IRB ID: 13-416

Approval Date:	9/7/2016	**Date for Continuing Review:**	10/13/2017
Submission Type:	Modification	**Review Type:**	Expedited

The project referenced above has received approval from the Institutional Review Board (IRB) at Iowa State University according to the dates shown above. Please refer to the IRB ID number shown above in all correspondence regarding this study.

To ensure compliance with federal regulations (45 CFR 46 & 21 CFR 56), please be sure to:

- **Use only the approved study materials** in your research, **including the recruitment materials and informed consent documents that have the IRB approval stamp.**
- **Retain signed informed consent documents for 3 years after the close of the study**, when documented consent is required.
- **Obtain IRB approval prior to implementing any changes** to the study by submitting a Modification Form for Non-Exempt Research or Amendment for Personnel Changes form, as necessary.
- **Immediately inform the IRB of (1) all serious and/or unexpected adverse experiences** involving risks to subjects or others; and (2) **any other unanticipated problems involving risks** to subjects or others.
- **Stop all research activity if IRB approval lapses**, unless continuation is necessary to prevent harm to research participants. Research activity can resume once IRB approval is reestablished.
- **Complete a new continuing review form** at least three to four weeks prior to the **date for continuing review** as noted above to provide sufficient time for the IRB to review and approve continuation of the study. We will send a courtesy reminder as this date approaches.

Please be aware that IRB approval means that you have met the requirements of federal regulations and ISU policies governing human subjects research. **Approval from other entities may also be needed.** For example, access to data from private records (e.g. student, medical, or employment records, etc.) that are protected by FERPA, HIPAA, or other confidentiality policies requires permission from the holders of those records. Similarly, for research conducted in institutions other than ISU (e.g., schools, other colleges or universities, medical facilities, companies, etc.), investigators must obtain permission from the institution(s) as required by their policies. **IRB approval in no way implies or guarantees that permission from these other entities will be granted.**

Upon completion of the project, please submit a Project Closure Form to the Office for Responsible Research, 202 Kingland, to officially close the project.

Please don't hesitate to contact us if you have questions or concerns at 515-294-4566 or IRB@iastate.edu.