EFFECT OF MULTIPLE FLUENCY CUES ON FEELING OF KNOWING (FOK) JUDGMENTS

by

TASNUSA ENAM

IAN M. MCDONOUGH, COMMITTEE CHAIR
SHEILA R. BLACK
GRAHAM MCDougALL

A THESIS

Submitted in partial fulfillment of the requirements for the degree of Master of Arts in the Department of Psychology in the Graduate School of The University of Alabama

TUSCALOOSA, ALABAMA

2018
ABSTRACT

Metamemory is the process of inspecting and judging our own memory. One type of judgment is feeling of knowing (FOK) judgments that are similar to making self-evaluations during learning for later performance. These judgments are influenced by various kinds of fluency cues. Perceptual fluency is the ease of processing perceptual information and does not predict later memory performance while conceptual fluency is the ease of processing conceptual information and accurately predict later memory performance. In academic settings, if students learn to distinguish between accurate and inaccurate cues, they can use the knowledge to use accurate cues effectively for future performance. Past studies investigated the effect of single fluency cue on JOLs. In reality, however, making self-evaluation takes time like FOK judgments paradigm and unlike JOL judgments paradigm that may not be generalizable. Students are also influenced by multiple fluency cues. The present study investigated how multiple fluency cues, font size (large versus small) and level of processing information (deep versus shallow), influence FOK judgments. Seventy eight younger adults studied word pairs in large or small font size and were either directed for deep or shallow level of processing information. Results revealed a significant interaction such that font size was only accounted for when information was processed at a shallow level. Furthermore, participants were overall underconfident in their predictions for their memory performance. Results are discussed in terms of applications in everyday learning, making decisions and in academic settings.
DEDICATION

This thesis is dedicated to all the people who helped me to make my research possible including my parents, for always believing in me and giving me confidence in every step along the way, my sister, for always inspiring me and helping me focus on my priorities, and Dr. Ian McDonough, who constantly guided me with his valuable input to help me create and shape this project, and to all the generous people who donated their time to make this research possible.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>df</td>
<td>Degrees of freedom: number of values free to vary after certain restrictions have been placed on the data</td>
</tr>
<tr>
<td>F</td>
<td>Fisher’s F ratio: A ratio of two variances</td>
</tr>
<tr>
<td>M</td>
<td>Mean: the sum of a set of measurements divided by the number of measurements in the set</td>
</tr>
<tr>
<td>SD</td>
<td>Standard Deviation: measures the amount of variability or dispersion for a subject set of data from the mean</td>
</tr>
<tr>
<td>SE</td>
<td>Standard Error: measures how far the sample mean of the data is likely to be from the true population mean.</td>
</tr>
<tr>
<td>p</td>
<td>Probability associated with the occurrence under the null hypothesis of a value as extreme as or more extreme than the observed value</td>
</tr>
<tr>
<td>r</td>
<td>Pearson product moment correlation</td>
</tr>
<tr>
<td>t</td>
<td>Computed value of t-test</td>
</tr>
<tr>
<td>&lt;</td>
<td>Less than</td>
</tr>
<tr>
<td>=</td>
<td>Equal to</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

I would like to thank Dr. Ian McDonough, the chair of the thesis committee for guiding me at every step of this project, from first to last and whose valuable suggestions/input made this project possible. I would also like to thank my other committee members, Dr. Sheila Black and Dr. Graham McDougall (presented alphabetically with last names) for their suggestions and support. Additionally, I would also like to thank Dr. Deborah Eakin, for her valuable suggestions and for introducing me to the program, ListChecker Pro 1.2 for my stimuli creation.

I would also like to thank the UA Subject Pool system for allowing me to recruit participants from the database.
CONTENTS

ABSTRACT ........................................................................................................................ ii
DEDICATION....................................................................................................................... iii
LIST OF ABBREVIATIONS AND SYMBOLS ................................................................ iv
ACKNOWLEDGMENTS ..................................................................................................... v
LIST OF TABLES ............................................................................................................... vii
LIST OF FIGURES ........................................................................................................... viii
CHAPTER 1 INTRODUCTION ......................................................................................... 1
CHAPTER 2 METHOD ..................................................................................................... 11
CHAPTER 3 RESULTS .................................................................................................... 15
CHAPTER 4 DISCUSSION .............................................................................................. 17
REFERENCES .................................................................................................................. 22
APPENDIX ....................................................................................................................... 31
LIST OF TABLES

1. Mean Scores for FOK and recognition memory…………………………………………………………………29
2. Mean Calibration Scores between FOK and recognition memory for FOK and LOP……..30
LIST OF FIGURES

1. Effect of Font Size and LOP on FOK judgment ratings ..................................................26

2. Mean FOK and recognition memory scores .................................................................27

3. Mean calibration scores between predicted and actual recall performance .......... 28
CHAPTER 1
INTRODUCTION

Metamemory is the process of monitoring one’s beliefs, knowledge, and mental processes to understand how one’s memory operates (Nelson & Narens, 1990). One application of metamemory is in academia where students need to evaluate their knowledge to know how much they have learned (Finn & Tauber, 2015; Koriat, 1997). In schools and colleges, students are given suggestions on how to study effectively for proper retention of knowledge and improve their performance. These suggestions are based on research from both memory, such as spacing out study sessions, and metamemory, such as asking self-questions for comprehension of the concepts when reading (Koriat, 1997). However, students often miscalibrate their predictions of how much they learned with their actual performance and often study methods/strategies fail to inform students on why these discrepancies occur and ways to avoid them. Students should recognize the cues that cause discrepancies between their metamemory and memory so that they can avoid them during studying. One of the factors that commonly result in the metamemory and memory discrepancy is fluency, the ease with which we process information (Craik & Lockhart 1972; Lanska, Olds & Westerman, 2014). Sometimes fluency cues that result in deeper processing of conceptual knowledge allow students to learn the information thoroughly and efficiently (Lanska et al., 2014, Metcalfe et al., 1993; Rhodes & Castle, 2008; Schwartz & Metcalfe, 1992). However, often students’ subjective experiences with fluency cues do not coincide with what they have actually learned (Kornell & Bjork, 2009; Koriat, 1997). This discrepancy leads to overconfidence in students. Overconfidence results in poor performance
such as thinking they have learned the information well when they did not (Bjork, Dunlosky & Kornell, 2013; Schommer, 1990). Students could also be under confident, such as thinking they have not learned the material well when they actually have, resulting in greater time cost and inefficiency during studying (Finn & Tauber, 2015). Additionally, during studying, students are influenced not only by a single fluency cue but multiple fluency cues that affect their metamemory and their actual knowledge. To understand how metamemory and memory miscalibration occur, two broad aspects need to be investigated—how fluency cues affect metamemory judgments and how it affects memory performance. In this current study, my primary analysis will be taking the initial step in understanding how multiple fluency cues affect metamemory. As one of the key factors in the current experiment, it is essential to understand, first, the fundamental process of metamemory.

In laboratory settings, metamemory is often studied by asking participants to make subjective judgments on their memory for word pairs at either encoding or retrieval stages of the memory process. One type of encoding-based judgments is judgments of learning (JOLs). JOLs assess how well participants believe that they will learn the items later on. They are made immediately following a presentation of a stimulus, such as a cue-target pair (cat-apple, where cat is the cue and apple is the target). Retrieval-based judgments include feeling of knowing (FOK) judgments, which assess how well participants believe that they will recognize the items later, immediately after a retrieval attempt. FOK judgments require participants to rate their beliefs based on the items they saw during encoding and how well they can retrieve at least partial information. It also predicts future recall. Hence, it is a unique judgment in the sense that it consists of both encoding and retrieval components and will be the class of metamemory judgments studied in the current experiment.
Although metamemory judgments usually align with memory performance, research has shown that certain cues lead to inaccurate metamemory judgments due to the influence of different factors (Castel, McCabe & Roediger, 2007; Dunlosky & Matvey, 2001; McDonough & Gallo 2012; Rhodes & Castel, 2008; Schwartz & Metcalfe, 1992). As mentioned earlier, one of the cues that mediate the underlying metamemory process is fluency—the speed and ease of processing information (Lanska et al., 2014). The influence of fluency on metamemory and memory differs based on the type of metamemory judgment—encoding-based judgment that depends on encoding phase of the memory or retrieval-based judgments that depends both on encoding and retrieval phase of the memory (Lanska et al., 2014; Lupker, Harbluk & Patrick, 1991; Metcalfe, Schwartz & Joaquim, 1993; Rhodes & Castel, 2008; 2009; Schwartz & Metcalfe, 1992).

**Subjective experiences of fluency affect metamemory judgments**

Fluency can act both as a diagnostic cue (metamemory judgment accurately predicts memory performance) and as a non-diagnostic cue (metamemory judgment inaccurately predicts memory performance) in learning situations (Koriat 2008; Kelley & Bjork, 2007). At encoding, for instance, Koriat (2008) asked participants to study a number of cue-target pairs, rate the richness of the information and how much thoughts they can bring to mind based on the pairs. Next, participants rated their confidence with encoding-based judgments (JOLs) in successive trials. At test trials, they were asked to identify the correct target in a forced-choice twoalternative question. The experiment consisted of several study-test trials. Cue-target pairs that were correctly identified in the first test trial were removed from the future study and test trials.
Result showed that items that were more accessible in mind (more fluent) in the previous study trials were correctly identified at both first and final test trials. Here, more accessible items that were rich in description could bring more information to mind and hence these items were more fluent (Koriat 2006; Koriat 2008). In addition, JOLs were higher for items that were more easily learned showing that it accurately predicted the advantage of easily learned items. Overall, the findings implicated that fluency can predict accurate performance in some situations.

However, fluency may not always be diagnostic of whether items have been learned properly at encoding. When used inaccurately to assess one’s learning, it can result in poor performance. Sometimes, items that are easy to recall can be learned in a short period of time. Participants’ evaluations of learning these items are also effective (Metcalf & Kornell, 2005, 2007; Thiede & Dunlosky, 1999). In contrast, items that are harder to recall need more time to learn so that they can be properly retained in memory. However, in some cases students do not give enough time in studying these items because they believe they have already learned the material (Kornell & Bjork 2007). For example, students prematurely stop studying for items when they “feel” like they know the to-be studied materials at the present time, which often makes them overconfident in their assessment of learned information (Kornell & Bjork, 2007). Instead of studying those items, they move on to focus on other items, implying, they stop their learning too soon. Learning strategies that result in rapid fluency at encoding occur due to less elaborate and more shallow processing of information, which may lead to a discrepancy between metamemory and actual memory. For instance, students usually disregard spaced study sessions (studying at spaced intervals) that enhance long-term memory performance, and prefer single study sessions (cramming) that results in rapid forgetting (Hintzman, 1974; Kornell, Castel, Eich
& Bjork, 2010; Kornell & Bjork, 2007). At encoding, single study sessions may increase fluency and gives a false perception that the items have been learned, causing the participants to be overconfident in their learning assessment, such as during exams. On the other hand, spaced session has greater time cost for increasing fluency of information, but it consists of elaborate processing that enhances learning, resulting in a good outcome during exams (Dempster, 1988; Kornell & Bjork, 2007; Kornell et al., 2010).

Similarly, at retrieval, fluency may act as a non-diagnostic cue. Even though fluency increases when we bring more information to mind, the relative difficulty in retrieving these items may reduce one’s confidence in one’s memory (Schwartz et al., 1991; Winkielman, Schwarz & Belli, 1998). Winkielman et al. (1998) asked some participants to recall 12 childhood events and others 4 events. Those recalling 12 events were underconfident in their memory ability compared to those who recalled 4 events. In another study, Kelley and Lindsay (1993) gave participants a list of words to study that consisted of names and places. Following the study phase, participants were given a general information test. Some words on the study list were correct answers for the questions on the test, while some were related, but not the actual answers for the test. The latter enhanced fluency (due to a relatedness effect) for retrieving the answers but not the accuracy. Results showed that for incorrect answers, the speed, frequency and confidence in producing those answers increased noticeably. It was also evident even when participants were informed that the list consisted of incorrect answers, implying that when fluency for retrieving incorrect answers increased, participants thought of them as correct answers. In a similar study, Kelly and Jacoby (1996) gave participants a list of anagrams to solve (e.g. deriving cinema from iceman). Some participants were given the answers while others were only given the words without the answers. Thus, participants who saw the solution completed the
anagram faster than those who did not. When they were asked to rate their confidence level in solving the puzzle, participants who solved it faster gave a high confidence rating compared to those who did not. In this study, participants’ speed of solving the anagrams was misattributed to the fact that it was “easy” to solve, not to the fact that they were shown the solution earlier. In all these cases, participants did not allow a deeper level of processing the information and relied on incorrect cues that enhanced their fluency but not their accuracy. Hence, their metamemory and memory were miscalibrated.

Discrepancies between metamemory and memory cause limitations in learning within academic settings. For example, students may be misinformed by fluency cues such as highlighted texts, larger font size, or a familiar heading of a chapter during studying. These perceptual cues might lead them to think that they will remember the information when in reality they will not remember in later performance (Finn & Tauber, 2015). One of the common types of fluency that causes the discrepancy between metamemory and memory performance is perceptual fluency, ease of processing perceptual information (Jacoby 1983; Whittlesea 1993; Lanska et al., 2014; Lupker et al., 1991; Rhodes & Castel, 2008; 2009). In contrast, conceptual fluency, the ease of processing conceptual information, has been shown to be an accurate predictor for memory performance, making it a relevant cue when making metamemory judgments (Metcalfe et al., 1993; Rhodes & Castle, 2008; Schwartz & Metcalfe, 1992). Perceptual fluency cues, such as font size, have practical applications in learning, such as fluency for reading textbooks. Hence, font size (large versus small font size) was used as one of the main manipulation in the study. Usually for conceptual fluency, one of the common cues used in previous studies was repetitive priming of the words (Metcalfe et al., 1993; Schwartz & Metcalfe, 1992). However, this kind of manipulation consists of both conceptual and perceptual
fluency aspects. For instance, seeing the items repeatedly can be both a perceptual cue (we perceive it repeatedly) and a conceptual cue (with repetitive perception, the meaning of the item is conceptualized). An example of conceptual fluency that may not have the influence of perceptual fluency is deep level of processing (LOP) versus shallow level of processing (LOP). This cue has been well-established conceptual cue in the literature (., et al 1991; Lanska et al., 2014) and was used as the manipulation for conceptual fluency in the current study.

Discrepancies resulting from perceptual fluency have been studied in metamemory judgments, with most studies focusing on the effect of a single fluency manipulation on JOLs or confidence judgments (Kelley & Lindsay, 1993; Kelley & Jacoby 1996; Koriat 2008; Mueller, Tauber & Dunlosky, 2012; Rhodes & Castel, 2008; 2009; Schwartz et al., 1991; Winkielman et al., 1998). Studies on how a single fluency manipulation affects FOK judgments also exist. For example, cue priming usually considered as a conceptual measure in earlier studies, increase FOK ratings without any significant effect on recall or recognition (Metcalf et al., 1993; Schwartz & Metcalfe, 1992). Since FOK judgments consist of both encoding and retrieval components, it is an ideal metamemory judgment to examine how both perceptual and conceptual fluency cues at encoding can influence later recall. The effect of fluency on FOK judgments can be extended in academia when students measure their beliefs of how much they learned after studying and self-testing themselves.

**Effect of multiple fluency cues on metamemory judgments**

As mentioned previously, most studies have looked at a single fluency manipulation and its effect on memory and metamemory judgments. In everyday learning situations, however, students may be subjected to more than one type of fluency cue. For example, when taking a test, students’ abilities to bring information to mind might be influenced by how easy it was to read
the texts in the book and whether they understood the concepts. Here, the fluency cues are font size and level of processing. However, studies of how multiple fluency cues affect metamemory judgments are few and are mainly focused on JOLs. Among those studies, Rhodes and Castel (2008) investigated how multiple fluency (relatedness and font size of word pairs) affect memory and metamemory judgments (JOLs). Participants were asked to study a list of paired associates (e.g., cat-apple). Results revealed that participants took both the diagnostic cue, in this case relatedness of the word pairs, and the non-diagnostic cue, in this case font size of the word pairs, into account. However, they gave different amounts of importance to the two cues when making their metamemory judgments. Specifically, mean JOL ratings were higher for related pairs ($M = 74.90, SE = 3.12$) than unrelated pairs ($M = 25.66, SE = 3.21$), $F(1, 27) = 163.99, \eta^2_p = .86$. Recall was also higher for related than unrelated pairs. Most importantly, the effect of font size remained for JOL ratings. The effect of font size for JOLs was smaller than the effect of relatedness ($\eta^2 = .17$ vs. $\eta^2 = .86$), and did not interact with the type of pair (related vs. unrelated). Additionally, large font size had higher JOL ratings ($M = 52.02, SE = 2.74$) than small font size ($M = 48.55, SE = 2.49$), $F (1, 27) = 5.58, \eta^2_p = .17$. Font size had no significant effect on recall. First, the results demonstrated that conceptual cues (relatedness) predicted accurate memory. Additionally, these cues were weighed more heavily in terms of importance when metamemory judgments were made. Second, the results showed that multiple types of fluency cues were weighed differently even though they did not influence each other. Overall, in presence of a more diagnostic cue for memory, people still take irrelevant cues into account during encoding when making metamemory judgments, even when these cues do not predict actual recall.

These findings prompt further research inquiries of how multiple fluency cues affect
FOK judgments. The present study investigated this question with two fluency manipulations—font size and level of processing. It is also acknowledged that answers to this question will be fairly preliminary, since there are different ways of manipulating fluency. This initial step is vital since it will provide further framework on understanding the whole picture of how discrepancies between metamemory and memory occurs in presence of multiple fluency cues.

**How do multiple fluency cues influence our judgment?**

Although research on fluency cues have shown that multiple cues can be weighed differently in terms of importance without an interaction between the cues, it should be noted that measures used for conceptual cues may have also consisted aspects of perceptual cue (Park et al., 2012; Rhodes & Castel, 2008). Level of processing used in the current study focus mainly on conceptual processing of information and hence, can provide to be a better memory predictor. Participants’ beliefs on certain relevant conceptual cues to memory performance may make them to completely disregard irrelevant perceptual cues. This may mean that there could be an *interaction* between the cue types on FOK judgments. Participants may intuitively judge deeper level of processing information as a more relevant cue for accurately predicting memory performance and disregard other irrelevant cues such as font size. The latter might only have an effect in shallow level of processing information.

Hence, my hypothesis was that there would be an interaction between conceptual cue, level of processing, and perceptual cue, font size, on FOK judgments. Specifically, in the presence of deep level of processing information, font size was predicted to have no effect on FOK ratings. In other words, conceptual cue would *suppress* irrelevant perceptual cue. I defined this hypothesis as the suppression effect hypothesis. Font size was predicted only to have an
effect when participants were processing the information at a shallower level and hence, taking other irrelevant cues into account as well.

Alternatively, keeping in line with previous studies, an alternative hypothesis was made that there would only be a main effect of perceptual and conceptual cue on FOK judgment without any significant interaction. This weighted additive effect predicted that font size and level of processing would both increase FOK ratings with different effect sizes since fluency generally has been shown to increase metamemory judgments. Specifically, conceptual cues were predicted to have a larger effect size than perceptual cues since conceptual cues were shown to be weighed more in terms of importance when making metamemory judgments (Rhodes & Castel, 2008; Schwartz and Metcalfe, 1992). I defined this alternative hypothesis as the additive effect hypothesis.

Looking at memory performance was not the primary concern of the present study since the primary interest was what kinds of cues people take into account for making predictions. Regardless, some preliminary hypothesis was made in the direction that font size would have little to no effect on recognition memory while level of processing would increase memory performance. In addition, a further question that was asked as an exploratory analysis was to look at the extent to which multiple fluency types influence the discrepancy between memory and FOK judgments.
CHAPTER 2

METHOD

Study Design

This was an experimental study within-group design. There were two independent variables with two levels in each variable: Font size (large 48 pts versus small 18 pts) and level of processing (deep processing versus shallow processing). Font size was the manipulation for perceptual fluency and level of processing was the manipulation for conceptual fluency. The dependent variable was FOK judgment ratings.

Statistical Design

A 2 (font size: large, 48 pts vs. small, 18 pts in Arial font) x 2 (LOP: deep vs. shallow) repeated measures ANOVA was conducted to see the effect of competing cues on FOK judgments.

An exploratory analysis (not a part of the main analysis) was also conducted to study how multiple fluency cues influence participants’ calibration of FOK judgments with their actual memory performance. Memory calibration was calculated by subtracting mean recognition memory scores from FOK ratings (mean FOK-mean recognition memory scores). Greater negative values would mean greater underconfidence and less negative values (more positive) would mean smaller underconfidence or well-calibrated
Participants

Seventy-eight participants ($M = 18-25$ years; 14 males, 64 females; 93% Caucasian, 5% African American, 1% more than one race) were recruited from University of Alabama, PY 101 Subject pool. All participants were freshmen and were compensated with course credit. Study was designed and implemented online through open source Collector program software (https://github.com/gikeymarcia/Collector).

Materials

Materials consisted of 480 unrelated cue-target word pairs. All individual words were nouns and were acquired from the database of word association norm program, ListChecker Pro 1.2 (Nelson, McEvoy & Schreiber, 1998, 2004; Eakin, 2010; www.eakinmemorylab.psychology.msstae.edu). All word-pairs were equated on normative factors such as set size (strength of (in) direct connections between and among associates of words/number of words), concreteness (the degree that the concept denoted by a word refers to a perceptible entity), connectivity (connections among associations of words) and resonance probability (connections from the associations of the words back to the word itself). Pairs were randomly divided into 4 lists each with 120 word pairs. Each list was divided into 5 blocks with 24 pairs such that there were 6 equal pairs of study conditions: large deep condition, large shallow condition, small deep condition and small shallow condition.

Cue Manipulations

Font Size: Fonts were presented in Arial font in black color with large font size presented in 48 pts and small font presented in 18 pts. The manipulation was based on previous study on multiple fluency effect on metamemory judgment (Rhodes & Castel, 2008).
Level of Processing: Level of processing was categorized as deep versus shallow conditions. In the deep condition, participants were asked, “Does the word on the right fit in the word on the left?” In the shallow condition, participants were asked, “Does the word pairs have the letter ‘e’ in it?”

Procedure

All items in all trials were presented in black capital letters in the center of a white background. Before the actual experiment, participants were presented with one round of practice condition so that they were familiarized with the experiment. All participants signed and received consent from at the beginning and debrief form at the end respectively.

Study Phase: Initially, participants viewed a word pair in either large or small font size for 2 seconds each. Next, if it was a deep condition trial they were asked, “Does the word on the right fit in the word on the left?” If it was a shallow condition trial, they were asked, “Does the word pair have the letter ‘e’ in it?” Participants were asked to choose either Yes/No as answer options. This phase was self-paced. On average participants took 4 seconds to answer the questions. Participants viewed a total of 120 word pairs in 5 blocks, each block consisting of 24 word pairs. All the conditions were randomized. There were a total of four counterbalances, each with a new list of word pairs. All blocks were randomized across trials. All counterbalances were randomized for participants.

Recall Phase: Participants were presented with the cue and were asked to retrieve the correct target by typing in the answer. The cues were presented in default font size of 32 pts. They were given 10 seconds to type the target name in a blank box. This phase lasted for 240 seconds.
**FOK Judgments Phase:** The cues were presented again to participants in a medium font size, and participants were asked to rate their FOK judgments (*How likely will you recognize this item on a later test?*) on a scale of 1-100 with 1 indicating “Not recognize at all” and 100 indicating “Definitely recognize”. This phase was self-paced. On average participants took 2 seconds to respond.

**Recognition Phase:** Finally, they moved onto the recognition phase where each cue was presented with the target and four additional lures. Participants were asked to select the correct target. This phase was self-paced. On average participants took 3 seconds to respond.

Inter stimulus interval between study phase, recall phase, FOK judgment phase and recognition phase was 4 milliseconds.
CHAPTER 3

RESULTS

Main Analysis

A 2 (font size: large, 48 pts vs. small, 18 pts) x 2 (LOP: deep vs. shallow) repeated measures ANOVA was conducted to see the effect of competing cues, conceptual versus perceptual cues, on FOK judgments. As shown in Figure 1, The results revealed a main effect of LOP, $F(1,77)=95.62, p < 0.001, \eta^2_p = 0.55$ such that deep LOP had a higher FOK rating ($M = 0.28$, $SD = 0.024$) than shallow LOP ($M = 0.19$, $SD = 0.020$). There was no main effect of font size, $F(1, 77) = 2.53, p = 0.12$. The results also revealed a significant interaction between level of processing and font size on FOK judgments, $F(1, 77) = 9.376, p = 0.003, \eta^2_p = 0.11$ supporting the suppression hypothesis. A paired sample t-test was conducted to understand the interaction. The results revealed that there was a significantly higher FOK rating for large font size ($M = 0.21$, $SD=0.18$) compared to small font size ($M = 0.18$, $SD = 0.18$), $t(77) = 4.07, p < .001$ only in shallow level of processing. Font size had no effect in deep condition, $t(77) = 0.82, p = 0.42$. In other words, irrelevant perceptual cues, such as font size, only had an effect when participants did not take relevant conceptual cues, such as deeper level of processing information, into account. This finding was consistent with the suppression effect hypothesis made earlier.
Secondary Analysis

There was a main effect of LOP on recognition memory, $F(1, 77) = 103.58, p < 0.01$, $\eta^2_p = 0.57$. In addition, there was a marginal significant interaction between level of processing and font size on recognition memory, $F(1, 77) = 3.84, p > 0.05$, $\eta^2_p = 0.01$. Additional paired sample t-tests revealed the interaction may have been driven by font size in the shallow condition where effect of large font size was bigger than small font size on recognition memory $t(77) = 1.81, p < 0.07$. The marginal significance of the result were caused by multiple outliers in our study for recognition memory performance. When these outliers were taken out (total of nine participants), the results for font size in shallow condition was marginally significant $t(68) = 1.93, p = 0.058$. Font size had no effect in deep condition, with outliers, $t(77) = 1.03, p = 0.31$, and without outliers, $t(68) = 1.29, p = 0.20$.

Exploratory Analysis

As a part of the exploratory analysis, mean accuracy percentage of actual recall was calibrated against FOK ratings (Figure 3, Table 3). Mean calibration scores were calculated by subtracting FOK judgment ratings from scores of recognition memory. A 2 (LOP: deep vs. shallow) x 2 (large vs. small) repeated measures ANOVA revealed a significant main effect of LOP, $F(1, 77) = 12.89, p < 0.01$ such that participants were more underconfident in deep condition ($M = -0.34, SD = 0.20$) compared to shallow condition ($M = -0.30, SD = 0.18$). There was no significant interaction of font size or interaction between LOP and font size.
CHAPTER 4
DISCUSSION

The main idea of the present study was to investigate how multiple fluency cues affect FOK judgments. The goal was to understand what kinds of cues people take into account when making a decision for their future memory performance. Several new findings relay the importance of the current study.

First, previous studies have only studied effects of fluency cues on encoding based metamemory judgments such as JOLs. However, these results cannot be generalized to all metamemory judgments, such as retrieval based, FOK judgments since the judgments are different. While JOLs are influenced by encoding phase of the memory, FOKs are based on both encoding and retrieval stages of the memory (Nelson & Narrens, 1990). Second, findings from JOLs cannot also be generalized in real life learning situations. For example, when students in academic setting start reading their text books, they do not immediately start monitoring their memory and start making predictions on whether they will remember the items later on (JOL paradigm). Rather, after reading, there is an interval period between studying the items and selftesting and finally making predictions for their future memory (FOK paradigm). The difference in judgment types and their interaction with fluency cues are apparent in the current findings.
Specifically, contrary to previous studies which showed conceptual and perceptual fluency do not interact with metamemory judgments like JOLs, the current study revealed a significant interaction between fluency cues on FOK judgment ratings. It was found that large font size had a greater FOK rating than small font size only in shallow condition. Since the results found an interaction of fluency cues for FOK ratings, it was possible to rule out the weighted additive hypothesis which predicted only main effects of fluency cues without any interaction. Instead, results supported the suppression hypothesis which predicted that in presence of strong conceptual cues, irrelevant perceptual cues would have no influence on FOK rating. The interaction was driven by font size in shallow condition. From the interaction results, it can be implied that when participants were faced with conceptual cues such as deep level of processing information to make their predictions, they did not take font size into account, hence font size had no significant effect in deep condition. Furthermore, font size only had an effect when they were processing the word pairs at a shallow level implying that irrelevant cues seemed to be taken into account when information was not processed properly. The present findings of the interaction for FOK ratings provide valuable information in the metamemory literature such that conceptual cues suppress the influence of perceptual cues for making retrieval based judgments if materials is well understood by the participants. Otherwise, if information is processed at a shallow level, perceptual cues may heavily influence participants’ decisions for future memory performance. This is a different for encoding based JOLs where effect perceptual cues does not depend on how information is processed at a conceptual level.

Our secondary and exploratory analysis revealed that even though predictions for deep was greater than shallow level of processing, overall participants were underconfident in their memory performance. This underconfidence was driven mainly by level of processing
(conceptual fluency) and not font size (perceptual fluency). On average, participants were more underconfident in deep level processing compared to shallow level processing regardless of font size. This finding was also contrary to previous studies where the main idea was that first, participants should be less underconfident in deep versus shallow processing. This novel finding implies that even though participants took conceptual cues into account when making their FOK judgments, they did not have the knowledge to make judgements that these cues would help their accuracy for final memory performance. The other contrary finding from previous study was the interaction of font size on recognition memory in shallow condition, such that large font size had higher memory performance scores compared to small font size. This result may have stemmed from the fact that both font size and shallow condition were perceptual cues and it could have been the case that participants’ attention were more directed towards this dual perceptual cues and hence, resulting in an interaction in recognition memory.

The overall results of the present study can be applied in academic setting for students’ learning. First, it can be implied that when students make predictions of how much information was learned, they take relevant conceptual cues into account to make those predictions and disregard irrelevant perceptual cues. Specifically, if students understand the material at a deeper conceptual level, they will disregard irrelevant cues like font size for their predictions. Additionally, if they already have a shallow level of understanding of the material, then they will use irrelevant cues like font size to make quicker predictions, such as if they are studying the material in the last minute before exams. The average underconfidence found may indicate that students may take conceptual cues into account but they are not perfect at calibrating their final accuracy. Perhaps, they are not aware of the fact that conceptual cues actually enhance their memory accuracy. Hence, they will spend more time on materials instead allocating their time to
learning other items that have not yet been learned well. Their misassumption that greater study time would have greater recall is termed as the ‘labor in vain effect’ (Nelson & Leonesio, 1988) where participants inaccurately spend longer time for self-paced studying even when they do not need to. Their misassumptions make them underconfident in their memory performance. In fact studying items more or less does not have any significant effect on recall unless they have learned the material at a conceptual level. However, most participants stop their learning before they master the items well enough (Nelson & Leonesio, 1988). Applying to academics settings, it can be implied that sometimes students terminate their study sessions before they have learned it well. In addition, even when they learn it well, they do not know when to stop. Hence, suggestions given to students should include not only strategies on how to study effectively but also knowledge on what kind of cues actually improves their memory performance and indicators that can help them to know when to stop self-paced learning so they can use their time more efficiently.

The main question in the current study was how multiple fluency cues are taken into account when making a prediction, in this case making FOK judgments. Overall, our results showed people are more prone to take irrelevant cues into account when they are already processing the information at a shallower level compared to a deeper level of processing. Additional suggestions for improving learning can also emphasize the process of how irrelevant cues can misled students if they study in the last minute, so that students will be well informed and learn to avoid those cues.

Some of the limitations of the study include lack of diverse population since the sample size mainly include college-aged and Caucasian student population. Results might differ if diverse sample was used as it has been shown that people are more fluent and are overconfident
in their metamemory predictions for materials that are related to their own race (Hourihan et al., 2012). In addition, college aged participants may generally have more available resources on how to monitor their learning, what cues affect memory and hence better metamemory compared to participants with less educational background. Future studies should include diverse sample to generalize the findings. One confounding effect that was found in the current experiment was that perceptual fluency had an effect on recognition memory in shallow condition. As mentioned earlier this could be because both shallow and font size were perceptual cues. In shallow condition, there was a dual effect of perceptual cues and participants may have focused their attention on these dual cues, specifically larger font size as it may have been more salient than looking for letter “e” in word-pairs. This may have resulted in the interaction of cues types in recognition memory contrary to previous findings. Currently to address this issue, I am designing a study where the goal is to investigate whether this effect resulted from attentional focus of the participants on perceptual cues by presenting participants with shallow vs. neutral processing of information.

Information about how individuals process multiple fluency cues may be applied to a variety of situations such as how people make everyday decisions with multiple pieces of information. As shown from the results, people are generally underconfident in their calibration for memory performance with their metamemory judgments so the results extend ideas into designing interventions to help public be aware of cues which actually improve or have no effect on actual memory so they can be well calibrated with their predictions and their memory performance.
REFERENCES


Figure 1: Bar chart displaying a significant interaction between conceptual (LOP: deep versus shallow) and perceptual (Font size: 48 pts versus 18 pts) fluency cues on FOK judgment ratings, $F(1, 77) = 9.376, p = 0.003$.

FOK=Feeling of Knowing; LOP= Level of processing
Figure 2: Bar chart displaying the comparison between mean FOK judgment ratings and mean recognition scores. Overall, participants are underconfident in their predictions for memory performance.

FOK = Feeling of Knowing; RCG = Recognition
Figure 3: Bar graph displaying the mean calibration scores (mean FOK - mean recognition memory) between FOK judgment ratings and actual recognition memory. Greater negative values mean greater underconfidence and less negative values (more positive) mean smaller underconfidence or well-calibrated. Overall, participants were more underconfident in deep LOP compared to shallow LOP regardless of font size. 

FOK = Feeling of Knowing; LOP = Level of Processing
Table 1.

*Mean Scores for FOK and recognition memory*

<table>
<thead>
<tr>
<th></th>
<th>Deep 48pts</th>
<th>Deep 18pts</th>
<th>Shallow 48pts</th>
<th>Shallow 18pts</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOK</td>
<td>28.01</td>
<td>28.72</td>
<td>20.87</td>
<td>18.53</td>
</tr>
<tr>
<td>Mean Recognition</td>
<td>62.14</td>
<td>63.38</td>
<td>51.41</td>
<td>48.68</td>
</tr>
</tbody>
</table>
Table 2.

*Mean Calibration Scores between FOK and Recognition Memory for LOP and Font Size*

<table>
<thead>
<tr>
<th>Font Size</th>
<th>Deep</th>
<th>LOP</th>
<th>Shallow</th>
</tr>
</thead>
<tbody>
<tr>
<td>48 pts</td>
<td>-34.17</td>
<td>-30.55</td>
<td></td>
</tr>
<tr>
<td>18pts</td>
<td>-34.73</td>
<td>-30.10</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX

August 10, 2017

Ian McDonough, Ph.D.
Assistant Professor
Department of Psychology
College of Arts & Sciences
The University of Alabama
Box 870348

Re: IRB #15-OR-335-R2 “Individual Differences in Memory and Metamemory”

Dear Dr. McDonough:

The University of Alabama Institutional Review Board has granted approval for your renewal application. Your renewal application has been given expedited approval according to 45 CFR part 46. You have also been granted the requested waiver of documentation of informed consent. Approval has been given under expedited review category 7 as outlined below:

(7) Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies.

Your application will expire on August 9, 2018. If your research will continue beyond this date, complete the relevant portions of the IRB Renewal Application. If you wish to modify the application, complete the Modification of an Approved Protocol Form. Changes in this study cannot be initiated without IRB approval, except when necessary to eliminate apparent immediate hazards to participants. When the study closes, complete the appropriate portions of the IRB Study Closure Form.

Should you need to submit any further correspondence regarding this proposal, please include the above application number.

Good luck with your research.

Sincerely,
UNIVERSITY OF ALABAMA
INSTITUTIONAL REVIEW BOARD FOR THE PROTECTION OF HUMAN SUBJECTS
REQUEST FOR APPROVAL OF RESEARCH INVOLVING HUMAN SUBJECTS

I. Identifying information

Principal Investigator: Ian McDonough
Second Investigator: Tasnava Enam
Third Investigator: [Redacted]

Department: Psychology
Department: Psychology
Department: Arts & Sciences

College: Arts & Sciences
College: Arts & Sciences

University: The University of Alabama
University: The University of Alabama

Address: 168 Gordon Palmer Hall
Address: 168 Gordon Palmer Hall

Telephone: (205) 348-7907
Telephone: (205) 348-7907

Fax: [Redacted]
Fax: [Redacted]

E-mail: [Redacted]
E-mail: [Redacted]

Title of Research Project: Individual Differences in Memory and Metamemory

Date Submitted: June 30, 2017
Funding Source: Psychology Department

Type of Proposal: Renewal

Please attach a renewal application

Please enter the original IRB # at the top of this page

II. NOTIFICATION OF IRB ACTION
(to be completed by IRB)

Type of Review: Full board

IRB Action: [Redacted]

Date: 8-9-18

Items approved: [Redacted]

[Redacted]
Informed Consent for a Non-Medical Study (UA)

Study title: Individual Differences in Memory and Metamemory

Principal Investigator: Ian McDonough, Professor, Dept. of Psychology

You are being asked to take part in a research study. Dr. Ian McDonough, who is a professor at The University of Alabama, is conducting the study.

What is this study about? What is the investigator trying to learn?
The purpose of this study is to learn more about people's memory and how they assess their memory.

Why have I been asked to be in this study?
You have been asked to participate in this study because you are part of the Psychology 101 subject pool, you are at least 18 years of age, and have near 20/20 vision or wear contacts.

What will I be asked to do in this study?
If you choose to take part, we will ask you to pay attention to words and pictures on a computer screen, and make simple judgments on those items such as the likelihood you will remember the word or picture. We will also be surveying your memory for the items that you viewed and have you make simple judgments about your memory such as the confidence in your memory for that word or picture. Additionally, we are going to have you fill out some questionnaires relevant to the task you just completed, your everyday memory performance, financial knowledge, personality, stress level, concussion history, sleep history, socioeconomic status, and demographic information.

How much time will I spend being this study?
We expect that it will take up to 1.5 hours of your time to complete the tasks and surveys.

Is the researcher being paid for this study?
This study is not being supported by any grant, agency or company.

Is this research developing a product that will be sold, and if so, will the investigator profit from it?
This research is not developing any product that will be commercialized afterwards.

How many people will be in this study?
About 1,000 other people will be in this study.

Will being in this study cost me anything?
There are no costs for participating in the study.

**What are the benefits that may happen if I am in this study?**
You will not directly benefit from participating, but your information will help us learn more about memory and why some people differ in their memory abilities.

**What are the benefits to science or society?**
The results of this study will help psychologists understand how people assess their own learning and will enable the groundwork for future studies to enhance the learning process.

**Will I be compensated for being in this study?**
You may earn 1.5 experimental credits per hour (up to 2 credits maximum), pro-rated in 0.5 credit increments based on your actual amount of participation.

**How will my privacy be protected?**
You may decide not to participate, choose not to answer any question, or stop participating at any time without any penalty.

**What are the risks (dangers or harms) to me if I am in this study?**
There are minimal risks involved in this study. You may feel tired, bored, or anxious from having your memory assessed.

**How will my confidentiality be protected?**
The records of this study will be stored securely and kept private. Authorized persons from The University of Alabama, members of the Institutional Review Board have the legal right to review your research records and will protect the confidentiality of those records to the extent permitted by the law.

**What are the alternatives to being in this study? Do I have other choices?**
The alternative to being in this study is not to participate. Remember that you can always participate in other experiments or complete a research paper that involves a similar amount of time and effort if you don’t want to participate in this particular study. Ask your PY101 TA for more information about these alternatives.

**What are my rights as a participant in this study?**
Taking part in this study is voluntary. It is your free choice. You can refuse to be in it at all. If you start the study, you can stop at any time. There will be no effect on your relations with The University of Alabama or the Department of Psychology.

The University of Alabama Institutional Review Board ("the IRB") is the committee that protects the rights of people in research studies. The IRB may review study records from time to time.
time to be sure that people in research studies are being treated fairly and that the study is being carried out as planned.

Who do I call if I have questions or problems?
If you have questions, concerns, or complaints about the study right now, please ask them. If you have questions, concerns, or complaints about the study later on, please email the laboratory at mae2research@ua.edu or call the investigator, Dr. Ian McDonough, at 205-348-1168.

If you have questions about your rights as a person in a research study, call Ms. Tanta Myles, the Research Compliance Officer of the University, at 205-348-8461 or toll-free at 1-877-820-3066.

You may also ask questions, make suggestions, or file complaints and concerns through the IRB Outreach website at http://osp.ua.edu/site/PRCO_Welcome.html or email the Research Compliance office at participantoutreach@bama.ua.edu.

After you participate, you are encouraged to complete the survey for research participants that is online at the outreach website or you may ask the investigator for a copy of it and mail it to the University Office for Research Compliance, Box 870127, 358 Rose Administration Building, Tuscaloosa, AL 35487-0127.

I have read this consent form. I have had a chance to ask questions. I am at least 18 years of age. I agree to take part in it.

I will receive a copy of this consent form to keep.

Signature of Research Participant

Date

Signature of Investigator

Date

35