EZGİ MELİSA YÜKSEL

THE EFFECT OF PERCEPTUAL FLUENCY ON ACCURATE, FALSE AND PREDICTED MEMORY PERFORMANCE

Bilkent University 2020

THE EFFECT OF PERCEPTUAL FLUENCY ON ACCURATE, FALSE AND PREDICTED MEMORY PERFORMANCE

A Master's Thesis

by

EZGİ MELİSA YÜKSEL

Department of Psychology

İhsan Doğramacı Bilkent University

Ankara

June 2020

THE EFFECT OF PERCEPTUAL FLUENCY ON ACCURATE, FALSE AND PREDICTED MEMORY PERFORMANCE

The Graduate School of Economics and Social Sciences
of
İhsan Doğramacı Bilkent University

by
EZGİ MELİSA YÜKSEL

In Partial Fulfillment of the Requirements for the Degree of
MASTER OF ARTS IN PSYCHOLOGY

THE DEPARTMENT OF
PSYCHOLOGY
IHSAN DOĞRAMACI BILKENT UNIVERSITY
ANKARA
June 2020

ABSTRACT

THE EFFECT OF PERCEPTUAL FLUENCY ON ACCURATE, FALSE AND PREDICTED MEMORY PERFORMANCE

Yüksel, Ezgi Melisa
M.A. in Psychology
Supervisor: Asst. Prof. Dr. Miri Besken

June 2020

The retrieval of memories does not reflect the exact copy of the original event and may include false information. Studies show that people become more susceptible to false memories due to post-event misinformation. One factor that might make the retrieval of the original event more problematic is the perceptual fluency of the information. If participants cannot clearly see the event, they might have an increased potential to integrate more false memories from the post-event knowledge. Finally, participants' predictions during encoding about how they will remember the original event might change, depending on the perceptual fluency, ease, and clarity of experiencing the original event. The current study aimed to examine the effects of perceptual fluency on accurate, false, and predicted memories. In three sets of experiments, participants were presented with picture stories, either in a fluent or disfluent form in a within-subjects design in the encoding phase. In the post-event

misinformation phase, participants saw all the stories that they saw in the encoding phase again in a fluent format, with some of the details changed. At the test phase, participants' actual and false memories were measured through a forced-choice recognition test with three-choices: correct, misinformation, and foil. In Experiment 1, participants were also asked to rate their future memory performance through Judgments of Learning (JOLs). In Experiments 2 and 3, JOLs were not collected to control JOLs' reactivity. Additionally, in Experiment 3, the possible effect of guess responses was controlled. The results of three experiments revealed that there was a consistency between predicted and actual memory for the disfluent items: participants' JOLs and memory performances were lower for the disfluent images than the fluent images. Participants showed a tendency to choose misinformation over the unrelated choice (i.e., foil), indicating that the misinformation manipulation increased the susceptibility to false memory. Contrary to predicted and actual memory, the disfluent or fluent presentation did not make any significant difference in the rate of false memories (susceptibility to misinformation). The results were in line with the perceptual fluency hypothesis and false memory literature.

Keywords: False Memory, JOLs, Memory Performance, Perceptual Fluency

ÖZET

GÖRSEL KANALLARDAKİ ALGISAL BOZUKLUKLARIN DOĞRU, SAHTE VE TAHMİN EDİLEN BELLEĞE ETKİSİ

Yüksel, Ezgi Melisa
Yüksek lisans, Psikoloji
Tez Danışmanı: Dr. Öğr. Üyesi Miri Besken
Haziran 2020

Anılarımız yaşadıklarımızın tam kopyası olarak çağrılmazlar ve yanlış bilgiler içerebilirler. Çalışmalar insanların olay sonrası maruz kalınan yanlış bilgilerden ve algısal akıcılık bozukluklarından dolayı sahte anı çağırmaya daha açık hale geldiklerini göstermiştir. Uyaranın görsel akıcılığı bellek hatalarının artmasına sebep olan faktörlerden biri olabilir. Bir başka deyişle, eğer katılımcılar bir olayı akıcı bir şekilde göremezlerse onların ikinci bilgiyi belleklerine entegre etme ihtimalleri artabilir. Ayrıca, katılımcıların kodlama süreçlerinde olayı ne kadar akıcı, ve kolay algıladıkları onların bu olay hakkındaki bellek yargılarını etkileyebilir. Güncel çalışma görsel akıcılığın gerçek, sahte ve bellek tahmini süreçlerine etkisini araştırmayı hedeflemektedir. Üç çalışma setinde katılımcılara bozulmamış (akıcı) ve maskelenmiş (akıcı olmayan) resimler aracılığıyla sekiz öykü sunulmuştur. Olay sonrası yanlış bilgi manipülasyonu kapsamında katılımcılara bu sekiz öyküdeki on iki detay değiştirilerek öyküler yeniden sunulmuştur. Gerçek ve sahte anılar; doğru, yanlış ve alakasız seçeneklerden oluşan üç şıklı tanıma belleği testiyle ölçülmüştür.

Bunlara ek olarak, 1. deneyde katılımcılardan öykülerin detaylarıyla ilgili bellek tahminlerini 0'dan 100'e kadar olacak şekilde belirtmeleri istenmiştir. Bellek yargısı bildirmenin etkisini kontrol etmek amacıyla 2. ve 3. deneylerde bellek yargısı toplanmamıştır. Ayrıca, 3. deneyde tahmin edilerek verilen cevapların etkisi kontrol edilmiştir. Bu 3 deney setinde bellek tahminleri ve gerçek bellek performansı arasında tutarlılık bulunmuştur. Bir başka deyişle, katılımcılar akıcı olmayan uyaranlar için düşük bellek performansı tahmin etmişler ve bu uyaranları daha az hatırlamışlardır. Ayrıca, katılımcılar yanlış seçenekleri alakasız olan şıkka kıyasla daha fazla seçmişlerdir. Bu durum yanlış bilgi manipülasyonunun sahte anı oluşturduğunu göstermektedir. Tahmin edilen ve gerçek bellek performanslarının sonuçlarının aksine, akıcı ve akıcı olmayan sunumlar arasındaki fark sahte anılarda görülmemiştir. Şöyle ki, katılımcıların sahte anıları akıcı ve akıcı olmadan sunulan uyaranlar için benzer oranlardadır. Bu sonuçlar literatürdeki sahte anı ve algısal akıcılık literatürü ile paralellik göstermektedir.

Anahtar kelimeler: Algısal Akıcılık, Bellek Performansı, Sahte Anı, Üstbellek, Öğrenme Yargıları

ACKNOWLEDGEMENT

Firstly, I would like to express my sincere gratitude to my advisor Asst. Prof. Dr. Miri Besken for her valuable guidance and support. I appreciate her patience, understanding, and the time she spent for me in the process of MA. I have learned a lot from her advice and our discussions. I would also like to express my gratitude to Asst. Prof. Dr. Aslı Kılıç Özhan for her guidance and warm attitudes since my undergraduate times and being a part of my examining committee. I would like to thank Assoc. Prof. Dr. Hüseyin Boyacı for his mind-opening questions during his enjoyable class and being a member of the examining committee.

I would also like to express my gratitude to the Scientific and Technological Research Council of Turkey (TUBITAK) for their financial support during my MA process under the program of BIDEB 2210-A.

I owe special thanks to my dearest family (Gülistan, Hasan Ali, Ulaş, Asya, Durmuş, and Gülşen) for their endless support. My brother, Ulaş, is always being there whenever I need in my life journey. I also owe special thanks to Hatice Dedetaş for our shared dreams, experiences, being a partner in academic life, her warm, and guidance in every part of my life; Cennet Süzme for being part of my growth, her perspective, and unconditional love; Melis Odabaş for her perspective-opening

lifestyle, and our crazy experiences; Gözde Sarı for each moment we shared and being my joy in life and my friends: Uğur Şatır, Cem Tosun, E. Hazal Ceylan, Alım Yılmaz, Merve Buçakcı for being part of my great times in Ankara. My journey could not be so wonderful without them.

I owe my appreciation for the valuable feedback, help, and emotional support offered by Gamze Nur Eroğlu. Thanks to our thesis-dates, the process of writing was full of fun and laughter. I am also grateful to Toygun Başaklar for his support in preparing my materials, Gizem Filiz, E. Eylül Ardıç, E. Cemre Solmaz for their valuable comments, and E. Hilal Kurt and other lab members who helped in the process of data collection for their time and support. I am also greatly thankful to all the participants for their valuable contributions, time, and patience.

TABLE OF CONTENTS

ABSTRACT vii

ÖZET ix

ACKNOWLEDGEMENT xi

TABLE OF CONTENTS xiii

LIST OF TABLES xvii

LIST OF FIGURES xviii

CHAPTER 1 1

INTRODUCTION 1

1.1 Reconstructive Memory 2

1.2 False Memory 3

1.2.2 Misinformation Manipulation 5

1.3 Metamemory & Fluency 11

1.4 The Effects of Perceptual Fluency on Predicted and Actual Memory 14

1.5 Perceptual Fluency, JOLs, and Susceptibility to False Memory 18

1.6 Main Questions of the Current Project 20

CHAPTER 2 23

EXPERIMENT 1 23

2.1 Method 24

2.1.1 Materials 24

2.1.2 Pilot Study and Power Analysis 26

2.1.3 Participants 28

2.1.4 Procedure 28

2.2 Results 32

2.2.1 Memory Accuracy 33

2.2.2 False Memory 35

2.2.3 JOLs 35

2.3 Discussion 40

CHAPTER 3 43

EXPERIMENT 2 43

3.1 Method 43

3.1.1 Participants 43

3.1.2 Procedure 45

3.2 Results 45

3.2.1 Accuracy 45

3.2.2 False Memory 47

3.3 Discussion 47

CHAPTER 4 50

EXPERIMENT 3 50

4.1 Method 51
4.1.1 Participants 51
4.1.2 Materials 51
4.1.3 Procedure 52
4.2 Results 53
4.2.1 Without Box Control 53
4.2.1.1 Accuracy 54
4.2.1.2 False Memory 55
4.2.2 With Box Control 56
4.2.2.1 Accuracy 57
4.2.2.2 False Memory 59
4.3 Discussion 61
CHAPTER 5 62
GENERAL DISCUSSION 62
5.1 Interpretation of Results and Theoretical Explanations 62
5.2 Limitations and Further Studies 71
5.3 Conclusion 74
REFERENCES 76
APPENDICES 88
APPENDIX A EXAMPLES OF CHANGES IN THE IMAGES 88
APPENDIX B FORMULAS FOR ACCURATE AND FALSE MEMORIES 89

APPENDIX C ORDER OF CONDITIONS – PARTIAL APPLICATION OF LATIN SQUARE COUNTERBALANCE 91

APPENDIX D COMPUTER SCREEN FOR FIRST INSTRUCTION 92

APPENDIX E COMPUTER SCREEN VIEW OF AN IMAGE AND JOLs 93

APPENDIX F COMPUTER SCREEN FOR RESTUDY PHASE INSTRUCTION 94

APPENDIX G COMPUTER SCREEN FOR TEST PHASE INSTRUCTION 95

LIST OF TABLES

Table 1 JOLs of Encoding Phase in Experiment 1 37

Table 2 JOLs of Misinformation Phase in Experiment 1..... 38

Table 3 Results of 2 (box control) x 2 (fluency) x 2 (types of response) ANOVA .. 60

Table 4 Means of Accurate and False Memories in Experiment 1, Experiment 2, and Experiment 3 60

LIST OF FIGURES

Figure 1 Four Main Stages of Experiment 1 25

Figure 2 Examples for Control Questions and Stimuli 31

Figure 3 Examples for Critical Questions and Stimuli 32

Figure 4 Four Main Stages of Experiment 2: Original Encoding Phase,

Misinformation Encoding Phase, Distractor Phase, Test Phase. 44

CHAPTER 1

INTRODUCTION

Imagine yourself witnessing a murder. You partially saw that the murderer used a knife. Then, the police officers came for interrogation, and they asked you whether the murder weapon was a gun. The misinformation that the weapon was a gun instead of a knife in this suggestive question might make you doubt yourself. Even further, later on, when you try to retrieve the original event in a subsequent interview, you might even misremember that you saw a gun instead of a knife.

The effects of post-event misinformation on later memory are well documented. Most of the time, when one receives a piece of suggestible information, they may mistakenly retrieve this post-event information as part of the original event. However, there might be factors as to how much this information is integrated into the original memory. One prominent factor that might change the level of integration of suggestible information may be the perceptual clarity of the originally-encoded event. For example, if one witnesses the murder behind a tree from where they cannot see clearly, this could potentially change the probability for the eyewitness to

accept the false claim of the presence of a gun more easily. Moreover, this type of physical obstruction in the eyewitness interview may also change how one predicts their memory to be for critical information. The current thesis project aims to investigate the effects of perceptual fluency on memory predictions, actual memory performance as well as susceptibility to misinformation when the information is presented clearly or in a blurry manner.

1.1 Reconstructive Memory

Since we are constantly exposed to thousands of stimuli within the experience of daily events, it is possible that we do not recall all of them. Sometimes, we might use our semantic memory (schemas, scripts, pragmatic inferences) or some information presented to us after the event to complete our representations of past events (Bartlett, 1932; Loftus, 1979). Even if most of the time memories are recalled accurately, these various sources of events and information might produce distortions such as the substitution of information in the memory with outside information, omission of specific detail from memory, and addition of new details to the original memory (Bartlett, 1932).

Bartlett (1932) was the first introducer of the idea that recalling past events is a reconstructive process that is rife with error; thus, retrieval of memories does not reflect the exact copy of the original memories. In his most famous experiment, he made Edwardian English participants listen to a folk tale of Native American culture, War of the Ghosts. He drowned on the serial reproduction and the repeated reproduction techniques to demonstrate the reconstructive nature of memory. In the serial reproduction task, participants studied the material and then were asked to

recall the material. Subsequently, other participants studied the recall of the previous participants. In the repeated reproduction task, participants listened to the material. Then, they recalled the same material over and over again with intervals between each recall. In both tasks, Bartlett (1932) found that the initial material was changed in each recall session regardless of whether the recaller was the same person or not. As expected, the amount of distortion in the material was higher for the serial reproduction task than the repeated reproduction since the different people added or omitted the information in different ways. The interesting point was that participants changed stories according to their cultural schemas. Some participants, for instance, remembered canoes in the story as boats. Bartlett (1932) introduced cultural schemas as one of the factors which guided the construction process. These findings indicate the malleability of retrieving memories, which is the main interest of the current thesis.

1.2 False Memory

During the mid-1970s, the focus on the reconstructive nature of memory increased because of two critical reasons. First, there was critical evidence for the weakness of the idea that our recall is the exact copy of the memories. Second, memory errors produced problems in criminal cases. Several experiments supported the malleability of eyewitness (Schaaf, Alexander, & Goodman, 2008; Schooler, Gerhard, & Loftus, 1986) due to over-questioning and exposure to the misinformation after the event.

Since the 1900s, the distortions in the memory were observed through several memory tasks: sentence reproduction, recall tasks affected by prior knowledge, recall tasks affected by self-sorting during the encoding (Bergman & Roediger, 1999;

Hemmer & Steyvers, 2009; James, Thompson, & Baldwin, 1973). Besides the repeated reproduction and prior knowledge, expectations, feelings, suggested ideas from outside, interpretations, and outcome desires influence how the memories are retrieved throughout the reconstruction process (Gallo, 2013; Tyng, Amin, Saad, & Malik, 2017). The distorted version of the retrieved information is named as false memory. In many false memory studies, the experimenters have commonly elicited false memories by using the participants' prior knowledge in Deese-Roediger-McDermott (DRM) paradigm or by exposing them to the subsequent conflicting information in the misinformation paradigm. The focus of the current thesis project is the misinformation paradigm; however, it is important to introduce both the DRM paradigm and the misinformation paradigm to understand how false memories occur in different manipulations.

1.2.1 Deese-Roediger-McDermott (DRM) paradigm

Deese (1959) had introduced a technique to examine the effect of associations on the extra-list intrusions, which is the type of memory error where people randomly recalled the items from the outside of the list. He created a list of 12 words (e.g., dream, awake, snooze, bed), which were highly associated with the critical lure word sleep. He used the word association norms from Russell & Jenkins (1954) cited in Deese (1959). He presented participants with a list of words without the critical lure word and then asked them to recall as many words as possible. The critical lure word sleep was recalled at the same frequency as other presented words.

As he argued, he found that the extra-list intrusions were related to the strength of the associations between the items. Roediger and McDermott (1995) conducted two

experiments with the model of Deese (1959) to replicate the findings and to reach a further understanding of the remembering process. In their first experiment, they used the same procedure with Deese (1959), and then found 40% recall of the lure word, sleep, and a high confidence rate (3.3 out of 4). In their second experiment, they extended the number of words to twenty-four words and created new lists. They replicated the findings of Deese (1959) and their first experiment. These two papers demonstrated how prior knowledge and schemas might affect the recall process of the memories. After Roediger and McDermott's (1995) paper, the paradigm where false memories have been elicited by participants' prior schemas and associations between the items was called the DRM paradigm.

1.2.2 Misinformation Manipulation

Another way to alter memory is the misinformation paradigm. The misinformation paradigm aims to alter the recall of an event by giving post-event misinformation to the participants. In the misinformation paradigm, people witness an event, followed by post-event misinformation. Post-event misinformation can be delivered through various methods such as usage of suggestive questions (e.g., Loftus & Palmer, 1974), questions with false information (e.g., Loftus, 1975), false stories about childhood memories (e.g., Loftus & Pickrell, 1995), narrative including misinformation (e.g., Zhu et al., 2010), repeated retrieval (e.g., Roediger III, Jacoby, & McDermott, 1996), images with inaccurate information (Okado & Stark, 2005; Stark, Okado, & Loftus, 2010), etc. Finally, the participants are asked to retrieve the events that happened in the encoding phase of the experiment.

In one of the earliest examples of the misinformation paradigm, Loftus and Palmer (1974) conducted an experiment to support the idea that the suggestive questions

influenced the reconstructive process. They presented a car accident video to the participants, and then they asked, “About how fast the cars were going when they smashed into each other?” and “Did you see any broken glass?” to participants. They used the words “bumped”, “collided”, “hit”, or “contacted” instead of “smashed” in different conditions to examine whether the meaning of the words would change the memories. As they expected, participants who were presented with the word “smashed” were more likely to predict the higher speed and more likely to recall broken glass than participants who heard the word “hit”. This study showed that the wording of post-event questions might change the content of the original event.

In another experiment, participants witnessed that a green car had an accident, and then they were exposed to the set of questions related to the car accident (Loftus, 1977). Since they aimed to create post-event misinformation, half of the participants responded a critical question, including misleading information, “Did the blue car that drove past the accident have a ski rack on the roof?” In the memory test, they selected the color of the car from a color scale. Participants in the critical question condition were more likely to choose blue or green-blue shades of colors.

To observe whether the findings of the misinformation paradigm were observed with different events, Loftus (1975) used movies, including different events (e.g., car accident). After watching the event, the participants answered either a question with correct information (e.g., How fast was car A going when it turned right?) or a question with false information (e.g., How fast was car A going when it ran the stop sign?). The higher percentage of participants in the stop sign conditions (53%) assumed they saw a stop sign during the movie compared to the turned right condition (35%). However, when they directly asked whether the participants saw a stop sign, they did not find many alterations in the memory. Loftus (1975) found that

the time of the misinformation was critical in the process of creating false memories. Participants' memories changed more with the misinformation before the memory test compared to the misinformation during the memory test. Since the results changed across the experiments, these findings raised critical questions about how suggestions might distort the memory in different settings.

Weinstein and Shanks (2010) also presented the list of objects' pictures to the participants. After studying the list of pictures, participants studied a list of objects' names, including unseen objects. In the recognition memory test, participants falsely assumed that they saw the unseen objects. In other words, their false alarm rates were high for the objects presented after the list only as names.

Different from the previous experiments, Okado and Stark (2005) created a misinformation paradigm, including images with misinformation, because they needed to present the original encoding phase and the misinformation phase in the fMRI. They examined whether the encoding processes of original and misinformation events were represented through neural activity or not. They presented participants with eight vignettes with fifty images in the original encoding phase. Subsequently, in the misinformation phase, participants restudied the same eight vignettes, each vignette including twelve changes from the original event. During both the original encoding and misinformation phase, Okado and Stark collected data for neural activities via fMRI. They tested participants' memories for two days later. In the test phase, they used multiple-choice recognition test. The options consisted of correct response (information from the original phase), foil response (information from the outside), and false response (information from the misinformation phase) for the details that changed. After the test, they asked

participants to report whether they saw the answer in the original phase, misinformation phase, or both.

They only analyzed the responses when the participants reported they saw the detail in either the original phase or in both phases. The results revealed that their method created a significant amount of false memories (47%). Since participants were not able to differentiate whether they received the information in the encoding phase or during the misinformation phase, Okado and Stark contended that their results could be explained by the source-monitoring hypothesis (Lindsay & Johnson, 1989). This hypothesis contends that the information from the post-event phase might sometimes be integrated with the original information, confusing as to where the source of information originates from. On a neural level, they found that higher neural activity in most of the brain regions could predict whether the participant would subsequently recall true or false memories.

Numerous studies show that misinformation can distort original memory traces. The distortive effects of misinformation on retrieving the original information are explained through various theories such as retroactive interference, fuzzy-trace theory, and source-monitoring framework. Retroactive interference theory claims that false memories are induced because participants update their memory traces when they re-experience the event or are suggested with misinformation in the post-event phase, leading to an irreversible change in the memory traces (Loftus & Pickrell, 1995). These updates delete the memory traces of the original information (Howe, 1998).

In another account, the fuzzy-trace theory defines forgetting as the disintegration of traces and the gradual fragmentation of memories (Reyna & Titcomb, 1997). As

retroactive interference theory, the fuzzy-trace theory also focuses on deletion, distortion, and omission of traces. Differently, the fuzzy-trace theory focuses on different versions of memory traces. The theory argues that memory traces consist of verbatim and nonverbatim (gist) information, generated after events occur or are imagined. Brainerd and Reyna (2005) found that the gist trace stayed longer, whereas the verbatim trace decayed quickly. They explained the presence of false memories with the idea that the post-event information distorts the memory trace of the original information, especially for the peripheral details (Loftus, Miller, & Burns, 1978).

Finally, the source-monitoring framework claims that false memories occur because people confuse the sources of original and post-event information (Johnson, Hashtroudi, & Lindsay, 1993; Lindsay & Johnson, 1989). Unlike the other hypotheses, the source-monitoring framework does not view false memory as the replacement of the initial information with the post-event information. Instead, the false memory is related to the misattribution of where the knowledge was learned. In other words, participants assume that they experienced the post-event information during the initial encoding phase. There are also brain-imaging studies, supporting the source-monitoring theory (Okado & Stark, 2005; Stark et al., 2010). These studies show higher activation for the subsequent false memories during the original phase and for the subsequent true memories during the misinformation phase (Stark et al., 2010). They attribute the strong activation to the robust encoding of the source information.

These theoretical explanations focus on various memory processes that explain the nature of the misinformation effect. Regardless of the nature of memory processes that lead to misinformation, some critical factors are prominent in changing the susceptibility to misinformation and false memory). These factors are also critical to

understand the scope of and reason for the misinformation effect. According to false memory literature, on the one hand, there are some situations in which people are more likely to be susceptible to misinformation: being under the age of 4 (Ackil & Zaragoza, 1998), aging (Schacter, Koutstaal, & Norman, 1997), being under high stress (Payne, Nadel, Allen, Thomas, & Jacobs, 2002), being distracted (Jacoby, Woloshyn, & Kelley, 1989), being under the time pressure (Benjamin & Craik, 2001), the long delay between event and recollection (Mudd & Govern, 2004), repeated exposure to misinformation and suggestion (Mitchell & Zaragoza, 1996) may increase false retrieval of information. On the other hand, there are some situations in which people are less likely to be susceptible to misinformation: exposure to the open-ended questions rather than leading questions, retrieving the information after a short rather than a long delay after the real event (Mudd & Govern, 2004), warnings (Gallo, Roberts, & Seamon, 1997), and being an adult (Ackil & Zaragoza, 1998) can reduce the rate of false memories.

The current thesis project examined the perceptual fluency, which is the subjective experience of ease while perceiving stimuli, as another factor which might influence the susceptibility to the misinformation effect for two reasons. Perceptual fluency might distort the encoding process since people cannot see stimuli clearly, which might increase memory errors. Perceptual fluency might also alter the control processes for encoding through metamemory, which might change the learning strategies (e.g., under-studying or over-studying fluent items). These reasons are discussed in more detail in the next section. In the literature, there were numbers of studies examining the effect of perceptual fluency on actual and predicted memory performances; however, it is not clear how perceptual fluency of item in encoding changes the effect of post-event misinformation. Before further discussing the

literature related to the effect of perceptual fluency on memory predictions and actual memories, it is critical to introduce the mechanisms of metamemory processes and their relationship with fluency. The next section introduces metamemory judgments in relation to fluency.

1.3 Metamemory & Fluency

Metacognition means cognition of our cognition, which consists of the processes of monitoring -evaluation of the cognitive processes- and controlling -regulation of cognitive processes- (Kuhn, 2000). One of the most studied metacognitive processes in the literature is metamemory, which means the memory of our memory. The initial definition of metamemory was "individual's knowledge of and awareness of memory, or anything pertinent to information storage and retrieval" (cited in Flavell and Wellman (1975), p. 6). Nowadays, metamemory refers to any heuristics, judgments, beliefs, and assumptions about the memory processes along with the regulation of the control processes for memory. Metamemory processes are critical to study because metamemory judgments regulate studying, learning, and remembering strategies (Goldsmith & Koriat, 2007; Son & Metcalfe, 2000).

Metamemory has been studied for several years through various methods such as the "feeling of knowing" (FOKs)," "retrospective judgments for the accuracy of the recall," and "judgments of learning (JOLs)" (Dunlosky & Bjork, 2008; Koriat & Goldsmith, 1996). One of the most common methods and the interest of the current thesis project, judgments of learning (JOLs), refer to the prediction of future memory performance in a subsequent memory test. In this specific method, participants are asked to make judgments about their learning during the encoding phase about how

confident they are about their subsequent memory performance. JOLs might be collected after either each item, called item-by-item JOLs or after the encoding phase, called aggregate JOLs. The current project uses item-by-item JOLs. In a typical experiment that assesses item-by-item JOLs, Rhodes and Castel (2008) asked participants to study a list of items, and after each item "to rate their confidence that they would later be able to recall that item on a scale from 0 (not confident at all) to 100 (very confident)" (p. 617). Typically, researchers seek whether these judgments are made in the same direction as the actual memory performance (resolution) or whether the judgments are predictive of the level of accuracy for the manipulation (calibration). (see Rhodes, 2016 for a detailed review).

Accuracy of JOLs was assessed based on the agreement between memory predictions and memory performance. In certain cases, JOLs can be quite accurate depending on the types of items, nature of the experiment, and so on (e.g., Mazzoni & Nelson, 1995; Dunlosky & Nelson, 1994). For instance, Tauber & Rhodes (2012a) found that participants' predicted and actual memory performances were higher for concrete compared to abstract items, indicating that participants were quite accurate in their judgments for subsequent memory performance for the comparison of abstract and concrete items (see a similar example in Hertzog, Dunlosky, Robinson, & Kidder, 2003). In a similar vein, JOLs were accurate for emotional words (Zimmerman & Kelley, 2010) and faces (Nomi, Rhodes, & Cleary, 2013) since there is consistency between memory performance and JOLs for emotional words and faces, both predictions and actual performances were high for emotional words and faces.

However, in other cases, despite using cues and heuristics, participants predict their future performance inaccurately, leading to metacognitive illusions (Koriat, 1997;

Koriat & Bjork, 2005). The metacognitive illusion for memory performance refers to the discrepancy between the predicted and the actual memory performance. Benjamin, Bjork, and Schwartz (1998) found, for example, participants assumed that they would more likely remember the answers to trivia questions that were easily retrieved from semantic memory at encoding would produce higher recall in a subsequent test; however, it was actually the items that were more difficult to retrieve at encoding that was remembered in a higher ratio.

Koriat (1997) introduced the cue-utilization framework as a method to interpret memory prediction success. The framework maintains that while people are making predictions on their memories, they evaluate the success of recall in a subsequent memory test by using inferences, rules, and heuristics instead of using the strength of their memories (Koriat, 1997). In other words, Koriat (1997) argued that while making metamemory judgments, people assessed the probability to recall or recognize the information based on a variety of extrinsic (i.e., cues related to conditions of learning -study time, number of items-), intrinsic (i.e., cues related to characteristics of the study items -imagery value of words-), and mnemonics cues (i.e., cues related to the subjective experience of processing -ease of processing-). There were some reported cues in the literature such as beliefs about task and manipulation (Mueller, Tauber, & Dunlosky, 2013), previous experiences (Hertzog et al., 1990), type of expected memory test (Mazzoni & Cornoldi, 1993), subjective experience of item difficulty (Arbuckle & Cuddy, 1969).

One factor that is shown to be highly effective in making memory judgments is fluency, despite the fact that fluency may not always affect actual recall in the same direction (Begg, Duft, Lalonde, Melnick, and Sanvito,1989; Shah and Oppenheimer, 2008). Oppenheimer (2008) defined fluency as “the subjective experience of ease or

difficulty associated with completing a mental task" (pg. 238). The effect of fluency on judgments was shown in a multitude of studies. Fluent information is typically perceived as more valid (Reber & Schwarz, 1999), more familiar (Wolk et al., 2004), and more likable (Reber, Winkielman, & Schwarz, 1998). Most importantly, for the current project, fluent processes and fluent information are typically perceived as more memorable (Besken & Mulligan, 2013; Mueller, Dunlosky, Tauber, & Rhodes, 2014). Thus, the relationship between the metamemory judgments and the actual memory performance is studied widely through JOLs and fluency (e.g., Besken & Mulligan, 2014; Koriat & Bjork, 2005; Matvey, Dunlosky, & Guttentag, 2001; Rhodes & Castel, 2008, 2009). The subjective experience of task difficulty -fluency of information- is manipulated through several factors such as encoding fluency (Hertzog, Dunlosky, Robinson, & Kidder, 2003), retrieval fluency (Koriat & Bjork, 2005), and perceptual fluency (Besken & Mulligan, 2013). The current project aims to investigate perceptual fluency as another factor which might alter the misinformation effect; therefore, it is critical to discuss how fluency influences actual and predicted (i.e., JOLs) memories. After introducing the metamemory judgment and its relationship with the fluency, the next section specifically introduces the effect of perceptual fluency on predicted and actual memory performance.

1.4 The Effects of Perceptual Fluency on Predicted and Actual Memory

The effect of perceptual fluency on predicted and actual memory has been a widely-studied topic over the last decade; however, the relationship between actual and predicted memory for perceptual fluency seems to change, depending on the type of perceptual fluency manipulation used. As a framework, the perceptual fluency

hypothesis suggests that easily and fluently perceived items produce higher memory predictions than items that are difficult to perceive, even though the ease of perception does not always play a role in the subsequent memory performance (Rhodes & Castel, 2008; 2009). There are several possibilities for how perceptual fluency affects memory performance and metamemory judgments.

In one case, perceptual fluency affects predicted memory performance but not the actual memory performance. Rhodes and Castel (2008) presented participants with the items in different font sizes (18 pt vs. 48 pt) to manipulate perceptual fluency. After each item, participants reported their confidence in recalling the item in a subsequent memory test on a scale ranging from 1 (not confident at all) to 100 (very confident). They found that people gave higher JOLs for the items in larger font size compared to the items in smaller font size; however, there was not any difference between the font sizes in terms of the actual memory performances. In this experiment, they observed the metacognitive illusion via perceptual fluency manipulation, meaning perceptual fluency might affect the JOLs but not the actual memory performance. In another experiment, Rhodes and Castel (2009) replicated their finding via auditory fluency manipulation. They presented participants with words in either quiet (conversational) volume or a loud volume. After each item, participants reported their likelihood of recalling accurately in a subsequent memory test. Even if there was no significant difference between the items in quiet and loud volume in terms of memory performances, participants gave higher JOLs for the items presented in loud volume than the items presented in quiet volume. In a similar line of work, Yue, Castel, and Bjork (2013) presented participants with clear and blurred words, followed by item-by-item JOLs. Their experiments showed a

significant difference between blurred or clear words in JOLs but not memory performance.

In another case, perceptual fluency may influence actual memory performance but not the predicted performance. Sungkhasettee, Friedman, and Castel (2011) showed 40 words to the participants in either inverted or rotated 180° or upright conditions. Participants were asked to report their JOLs after each item. They revealed that participants were more likely to recall the rotated items than the upright items; however, JOLs were not influenced by the rotation manipulation (see Besken, Solmaz, Karaca & Atılgan, 2019, for a similar result). This can be explained by the concept of desirable difficulties (Bjork, 1994): When participants have some difficulties during encoding because participants are putting more effort into encoding, this might produce deeper (Alter, Oppenheimer, Epley, & Eyre, 2007), and more careful encoding (Song & Schwarz, 2008).

In another possibility, perceptual fluency might affect JOLs and actual memory performance in opposite directions. In other words, a particular perceptual fluency/disfluency manipulation may increase memory performance while decreasing JOLs. Besken and Mulligan (2013) presented participants with intact and backward masked words and collected both aggregate (in experiment 1) and item-by-item JOLs (in experiment 2) for their prediction of future memory performance in the free-recall task. In both experiments, the results revealed that both aggregate and item-by-item JOLs were higher for the intact items than the masked items. However, their memory performance was higher for masked than intact words. Besken and Mulligan (2014) replicated the findings of the previous study with auditory disfluency manipulation. Participants were aurally presented with words, which were inter-spliced with silence or were intact. The results were consistent with the results

in the previous experiment. Both aggregate and item-by-item JOLs were higher for the intact items compared to the items inter-spliced with silence, with opposite effects on actual memory performance.

Even if most of the time, perceptual fluency manipulation ends up with inconsistencies between JOLs and actual memory performance, there are studies in which perceptual fluency manipulation produces consistency between predicted and actual memories. Yue, Castel, and Bjork (2013), for instance, examined the effect of perceptual fluency on actual and predicted memories through blurring manipulation. Participants studied blurred or clear words and gave JOLs after each item. They found partially accurate JOLs for blurred words, indicating that they are quite accurately predicting that they were not able to remember blurred words well. Moreover, Undorf and Zimdahl (2019) manipulated the perceptual fluency through font sizes (18 pt vs. 48 pt) to investigate its effect on JOLs and actual memory performance. They found an increase in both memory predictions and memory performance for larger font sizes compared to smaller font sizes; however, the difference between fluent and disfluent items in terms of memory performances was not high as much as the difference between the conditions in terms of JOLs.

As stated above, the effect of perceptual fluency on JOLs and memory might change depending on various factors. In the current project, the effect of perceptual fluency on memory is further studied in the misinformation paradigm to observe how it influences predicted and actual memory performance as well as susceptibility to misinformation.

1.5 Perceptual Fluency, JOLs, and Susceptibility to False Memory

Since the perceptual fluency manipulation influences JOLs, it is also critical to investigate how perceptual fluency manipulation influences false memories through metamemory judgments. Participants' JOLs are critical for the current thesis because the misinformation paradigm might change people's judgments (Frost, 2000). Thus, the perceptual fluency manipulation may change JOLs, which in turn might change the people's strategies to regulate studying and remembering (Dunlosky & Rawson, 2012; Nelson & Narens, 1990). Alter et al. (2007), for instance, found that people were more likely to choose more effortful and elaborative processing styles when they had lower confidence in their learning. Similarly, Hanczakowski, Zawadzka, and Cockcroft-McKay (2014) collected the people's feeling-of-knowing (FOKs) by asking participants to report how likely they would recognize the unrecalled item in the first cued-recall test. They found that participants were more likely to choose to restudy the items with high FOKs. Moreover, Undorf and Ackerman (2017) tried to examine whether JOLs prediction for the chance to recall in the future test- and study time were related. They found a U-shape relationship between the self-paced study time and JOLs, meaning that people were more likely to spend less time on the items with lowest and highest JOLs (from 1 to 100) while they were spending more time on the items with moderate level JOLs (between 21-60 out of 100), indicating that people change their strategies according to their JOLs.

As discussed above, people tend to improve their memories by using strategies such as restudying, longer study time, and so on. Since the disfluent presentation mostly produces a decrease in JOLs, people might show a tendency to put in more effort into encoding disfluent than fluent items by paying more attention or restudying to disfluent items. In contrast, because disfluent items are difficult to process,

participants may be less motivated to study those items less than fluent items, creating a different pattern of results for actual recall. Therefore, it was critical to investigate whether these strategies might be observed in the misinformation paradigm in which participants were presented with inaccurate information while restudying. Therefore, one of the aims was to investigate the influence of metamemory judgments on susceptibility to the misinformation effect.

Another focus was to examine the direct relationship between perceptual fluency and false memory. Since the disfluent presentation might distort the encoding process, memory errors might be higher for those. Recently, the relationship between perceptual fluency and false memory was examined through the DRM task (Sanchez & Naylor, 2018). Their main aim was to test two different ideas for the DRM task. On the one hand, some studies show that false memories might be reduced by increasing detailed item-based attention which will decrease the automatic processing (Hege & Dodson, 2004; Mccabe, Presmanes, Robertson, & Smith, 2004). On the other hand, other studies show that detailed item-based attention might increase false memories since the in-depth processing may increase the automatic spreading of activation (Loftus, 2004). Sanchez and Naylor (2018) decided to use perceptual disfluency to improve detailed item-based attention and observe whether the item-based focus would increase or decrease the number of false memories.

They used classical six DRM lists from Roediger and McDermott (1995), manipulating the disfluency in between subjects via font types (14 pt. Mistral font as disfluent and 14 pt. Arial font as fluent). They controlled the perceived task difficulty via scale from 1 to 10. Averages for two conditions were significantly different from each other. They used both recall and recognition as a final memory test. Even though they found a significant amount of false memory in fluent condition,

participants recalled and recognized the lure item more in the disfluent condition. However, they did not find any difference between fluent and disfluent conditions in the accuracy of the responses. They interpreted that the relationship between fluency and false memory was crucial to examine because even in the fluent conditions, they found a significant amount of false memory, which might be critical for situations involving eyewitness testimony.

In light of the previous literature, it is evident that metamemory judgments and fluency are critical factors in the misinformation paradigm; however, it is not clear how perceptual fluency changes the accurate, predicted, and false memory performance in the misinformation paradigm; and how perceptual disfluency of the original information influences the susceptibility to misinformation. Thus, the current thesis project examines how disfluent presentation may change participants' predictions, accurate and false memories in the misinformation paradigm and when people witnessed the initial event disfluently, whether their susceptibility to post-event misinformation increase due to their low JOLs, or not.

1.6 Main Questions of the Current Project

In the current project, participants were presented with stories consisting of images either in their intact form (i.e., fluently) or with a checker-board mask (i.e., disfluently). After the initial encoding phase, all the stories in the encoding phase were presented again in their intact form with twelve details' changes (e.g., the number of bananas that the girl bought changed from 2 to 3) as misinformation manipulation. Their memories for the details from the first encoding phase were tested through a forced-choice recognition test. Additionally, in Experiment 1,

participants rated their subsequent memory performance from a scale ranged from 0-100 (not confident I would remember quite confident I would remember, respectively) to examine how the metamemory processes would influence the relationship between fluency and false memory. Considering the related literature, and with this method, the current thesis tried to test some critical questions and hypotheses.

The first critical question was whether the perceptual fluency of the stimuli during encoding changed the amount of false and accurate memories in the final memory test. Depending on the question, the first hypothesis of the study was that the higher false memory would be observed in the disfluent presentation condition. This hypothesis was based on the idea that the disfluent presentation of the stimuli might make participants more susceptible to post-event misinformation. While investigating the distribution of the false memory to fluent and disfluent conditions, it was also critical to observe whether the fluency of the information during the first encoding would predict future memory accuracy. Since the disfluent presentation decreased the quality of the encoding process, the second hypothesis was disfluent presentation would significantly decrease the accuracy rates in the final memory test compared to the fluent presentation.

The second question was whether the fluency during the encoding phase influenced the people's JOLs. In the current study, it is expected that participants will give higher JOLs to fluent than disfluent items, in line with the most common findings in the literature (Besken, 2016; Besken & Mulligan, 2013, 2014; Rhodes & Castel, 2008, 2009). Moreover, since the metamemory judgment impacts the people's learning strategies as mentioned above, it was crucial to investigate the effect of JOLs in this relationship. The lower JOLs for the disfluent items would increase the

people's tendency to learn the following information more than fluent items, which may potentially lead to increased susceptibility to misinformation.

The third question was to explore whether the JOLs would differ between the initial encoding phase and the misinformation phase. To clarify, during the encoding phase, people were exposed to some events, and during the post-event information phase, they were exposed to the same events with small differences as misinformation manipulation. In both phases, they were asked to report their judgments for the future memory performance. Thus, it is crucial to investigate how the JOLs changed across different items (changed vs. unchanged items) and different times (initial learning vs. post-event information learning). There was not any particular hypothesis for the third question since it was not clear in the literature. The answers were explored for further studies.

CHAPTER 2

EXPERIMENT 1

In Experiment 1, it was to examine the effect of fluency on metamemory and false memory by using the misinformation paradigm. Participants were asked to study 8 stories, which presented through images. For half of the stories, the images were fluent, so they were presented in their intact form. For the other half of the stories, the images were disfluent, covered with a checkerboard pattern. Then they were exposed to misinformation about the stories by changing the visual details of the stories. All the images in the misinformation phase were presented in their intact form. During both the learning phase and the misinformation phase, participants' judgments of learning were collected to measure metamemory. In the end, participants answered multiple-choice questions about the stories. The main question was whether the fluency of the information during encoding changed the number of false memories. We also examined the question of how the predictions of the future memory performance alter the relationship between the fluency and the false

memory. The main hypothesis was people would have less accurate and predicted and more false memories for disfluent items compared to fluent items.

2.1 Method

2.1.1 Materials

The eight vignettes from Okado and Stark (2005) used to create false memories. Each vignette consisted of fifty images that flow as a story. The general context of the stories included both neutral events (e.g., a casual day in the school, shopping) and criminal events (e.g., fight, robbery). Okado and Stark (2005) made small changes in the misinformation phase of the study for the images *(see examples in Appendix A)* to create false memories. Each story had 12 critical changes that were presented in the misinformation encoding part. For instance, in the robbery vignette, the thief crushed the door of the car via a credit card in the original encoding phase, whereas he used a hanger in the misinformation phase. The images were used in both the initial encoding phase and the misinformation phase so that we could manipulate the fluency in the same manner during both encoding and misinformation phases. In other words, by using a similar type of stimuli, we could create a similar feeling of perceptual fluency in both the original encoding phase and the misinformation phase.

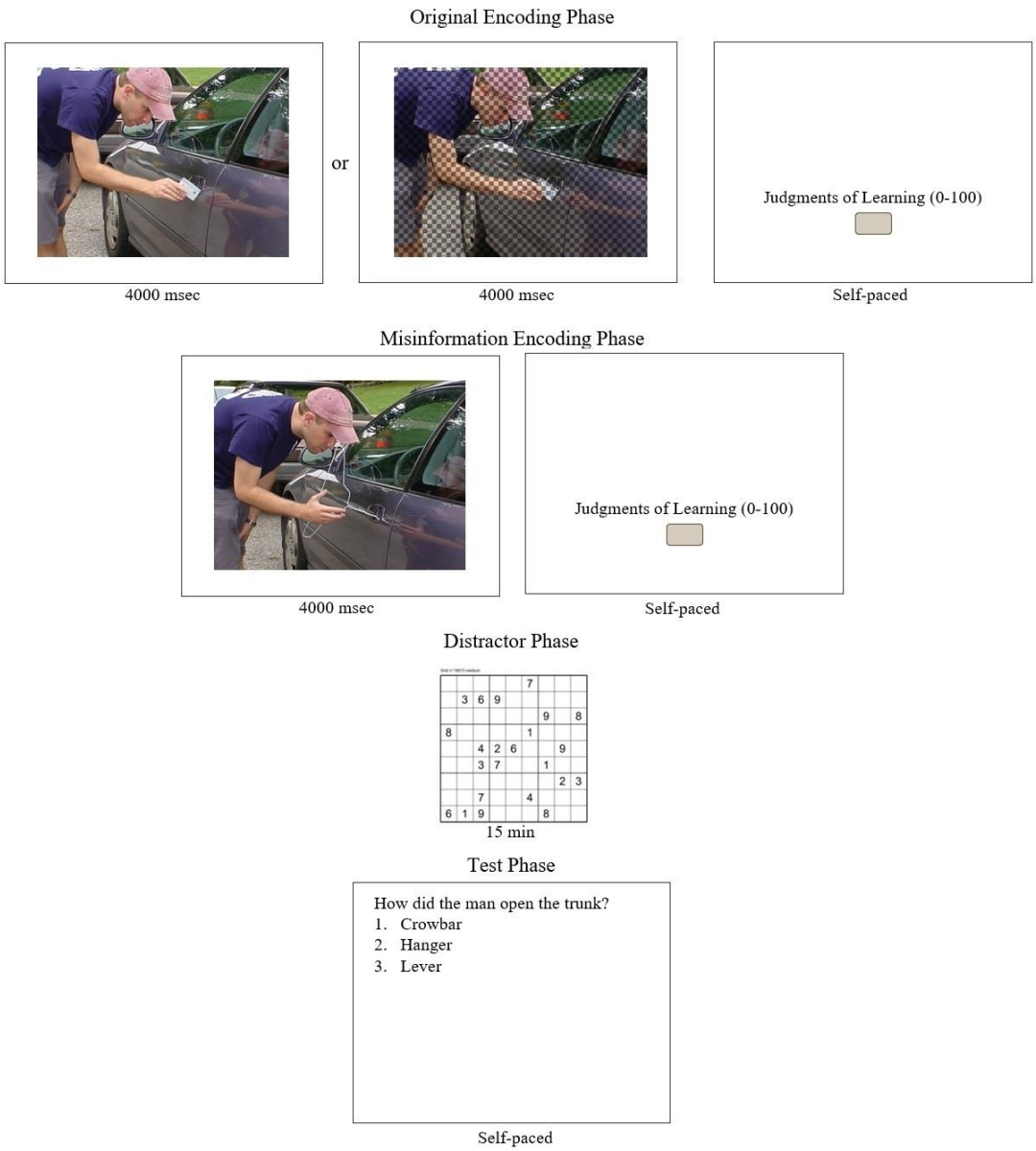


Figure 1 Four Main Stages of Experiment 1 Original Encoding Phase, Misinformation Phase, Distractor Phase, Test Phase. In the original encoding phase, participants either see the fluent or the disfluent image followed by JOLs. In the misinformation phase, participants see the fluent version of the image with a small change (e.g., credit card turns to a hanger). In the distractor phase, participants made a paper-based Sudoku Test. In the test phase, participants were presented with each question and asked to answer by using the numbers

The fluency of half of the stories (four out of eight) was distorted by using a checker-board mask *(see in Figure 1).* The checker-board mask was implemented to each image via MATLAB programming by using .05 intensity. The intensity of the mask was adjusted by asking pilot participants, whether they could see the specific items in the image, or not. However, it was crucial to be sure that participants would perceive the critical items due to the checker-board mask. Thus, after adjusting the intensity, the survey was conducted with twenty random images with critical items via Qualtrics. Participants were asked to report what was the critical object and to evaluate how much time they needed to understand the item. For example, they were asked to write the name of the object in a man's hand. The accuracy of the responses was higher than 90% for all the images. Thence, the images in disfluent conditions were visible.

Since there were eight different stories, the Latin-square counterbalance rule for the even-number of conditions was used for the order of stories. The second presentation rule was that each story was distributed into fluent and disfluent conditions equally. While adjusting the second role into the first one, the counterbalance rule for the stories was partially changed. At the end of the counterbalancing, there were eight different combinations of fluency and stories *(see in APPENDIX D – a)* The eight different conditions were prepared in E-Prime. Each condition had an equal number of participants who were assigned to conditions according to the order of participants.

2.1.2 Pilot Study and Power Analysis

There were some critical changes in the method and materials from the method of Okado and Stark (2005); therefore, we conducted a pilot study with thirty-two

participants (13 females, M_{age} = 21.8, SD = 2.17) from I. D. Bilkent University. Participants were given 20 Turkish Liras for their contributions. The pilot study aimed to control 15 minutes Sudoku test as a distractor in the misinformation paradigm and to determine the required participants' numbers through Power analysis.

A paired samples t-test with the accuracy percentage as dependent variable revealed that there was a significant effect for fluency, $t(31) = 4.71$, $p < .001$, $d = .82$. Fluent items ($M = .59$, $SD = .09$) produced higher memory performance than disfluent ones ($M = .52$, $SD = .1$). The other important hypothesis was related to the comparison of the fluent and disfluent conditions in terms of false memories. A paired samples t-test with the amount of false memories as dependent variable showed that there was no significant difference between the fluent ($M = .65$, $SD = .11$) and disfluent conditions ($M = .70$, $SD = .09$), $t(31) = 1.97$, $p = .058$, $d = .35$.

Depending on the results of the pilot study, a Post hoc power analysis was conducted through G*Power3 (Faul, Erdfelder, Lang, & Buchner, 2007) to test the difference between two dependent groups means using a one-tailed test, found effect size (=.35), and an alpha of .05. Results showed that with thirty-two participants, the study achieved a power of .62. Since the p-value was relatedly close to the .05, I made power-analysis to determine the required sample size to reach a power of .80. An a priori power analysis with a dependent group means using a one-tailed test, .35 effect size, and an alpha of .05 revealed that required participants' number as fifty-two. Because there were eight conditions, we decided to collect data from fifty-six participants.

2.1.3 Participants

Fifty-six students (34 females, M_{age} = 19.65, SD = 1.21) from I. D. Bilkent University participated in exchange for course credits. Participants were all native Turkish speakers and had a normal or corrected-to-normal vision. The ethical approval for the experiments was granted by Bilkent University Human Subjects Ethics Committee.

2.1.4 Procedure

Participants were invited to the laboratory to attend the experiments. After signing the informed consent, participants were given verbal instructions about the general procedure of the experiments. The exact nature and the aim of the experiments were not given to the participants to make the misinformation phase work. They were informed that the study aimed to examine whether the repeated study sessions strengthen the memory performance in a subsequent test and whether the strengths of the memory performance change depending on the fluency of encoding. Thus, participants were also warned that even if the general concepts of the stories would be the same in the second sessions, they needed to study carefully in the second encoding in order to see the effect of repeated study. The demographics information (age, gender, handedness) were collected via E-Prime and entered by the experimenters. After answering participants' further questions, the experiment began with a written instruction, which was similar to the verbal instruction *(see Appendix E)*.

The experiments consisted of four main sessions: original encoding, misinformation phase, distractor, and test (*see in Figure 1*). In the original encoding phase, participants were exposed to eight stories with fifty images. Half of the stories were

presented in a fluent condition with intact pictures, whereas the other half were presented in disfluent conditions with images covered with the checkerboard pattern. Each image was presented for 4000 ms with a 500-ms inter-trial interval (ITI). Okado and Stark (2003) used 3500 ms as a presentation time; however, the current study used disfluent images that required more time to be processed. In the pilot survey, besides asking participants to name the specific objects, they were also asked to choose how much time they needed to see the detail of the image from the options; 3 sec, 4 sec, and 5 sec. The average time was 4 sec. Thus, the 4000 ms was adjusted according to the results of the Qualtrics survey.

In the misinformation phase, participants were presented with the changed version of eight stories. All images were presented in their intact form, regardless of whether they were presented in fluent or disfluent forms during the initial encoding phase. Each story had 12 changed details to produce misinformation manipulation. Participants were told that the aim of the study was to see how multiple encodings of the same story may change memory performance without being revealed that some of the details had changed. During these two phases, participants were also informed about the end of each story, via written instructions after each story. After the instruction, participants decided when they moved to the next story by themselves to provide them with time to rest.

During these encoding and misinformation phases, participants were asked to pay attention to the details of the images. While studying each image, participants were asked to make item-by-item memory predictions about their future performance. After being presented with each picture, participants provided scores for how sure they are whether they would remember the item on a later memory test on a scale from 0 (not confident that I will remember this at all) to 100 (quite confident that I

will remember this on a later memory test). The experimenters warned the participants to use the whole scale while making their judgments. After encoding and misinformation phases were completed, participants were given Sudoku tests for fifteen minutes as a distractor task. Sudoku tests (medium level) were created on the www.printmysudoku.com website. Okado and Stark (2005) originally used a distractor period of two days to increase the time between encoding and test sessions. However, it was hard to schedule meetings with the same participants twice; therefore, we used 15 minutes of Sudoku as a distractor. In the pilot study, we controlled whether 15 minutes was long enough for misinformation manipulation to produce susceptibility to misinformation.

In the test session, participants completed a multiple-choice recognition memory test with 18 questions for each story (144 questions in total). The questions from Okado and Stark (2005) were used, and the questions were translated and adapted to Turkish. Twelve of the questions were related to 12 critical items (*see examples in Figure 3*). The order of the questions and the order of the choices were randomized across conditions. In each question of the recognition test, participants were presented with the information in the encoding phase, misinformation phase, and some new information (foil item) as options. They were asked to choose the correct answer in line with the details from the initial presentation of the event (encoding phase). In other words, the questions had three alternatives forced-choice (3AFC) recognition test: original item, misinformation item, and novel/foil item. The other six questions were control questions with two foil and one correct choice (*see examples in Figure 2*). All stories were tested in the same order they were presented during the encoding phase to hold the time difference between encoding and test phase equal for each story. In the study by Okado and Stark, participants also

completed a surprise source memory phase after the multiple-choice recognition phase. This phase was excluded from the current study since our research question was specifically related to the first presentation. Participants were asked to give their responses by thinking about the images in the encoding phase.

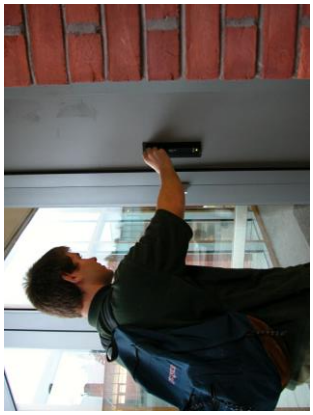
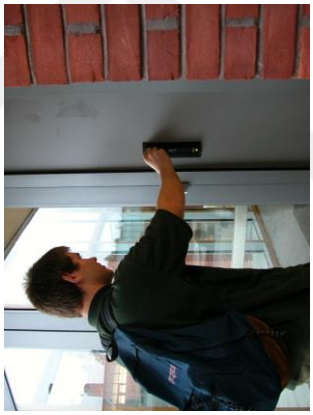
Control Questions	Misinformation Phase	Encoding Phase	Question
			How did the man open the trunk? 1. Crowbar **(foil)** 2. Hanger **(foil)** 3. Lever **(correct)**
			How did Sam get into the building? 1. Another student **(foil)** 2. Access card **(correct)** 3. It was not shown **(foil)**

Figure 2 Examples for Control Questions and Stimuli

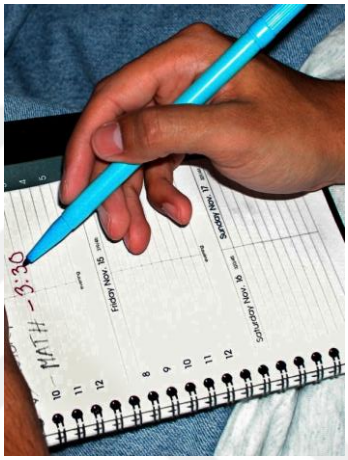
Critical Questions		
Encoding Phase	Misinformation Phase	Question
		What does Christina do to prevent Seth from watching the TV? 1. Close via TV remote **(correct)** 2. Close via the button **(mis)** 3. stands in front of it **(foil)**
		What time was Nicholas' next Math class going to be? 1. 9:00am **(foil)** 2. 1:15pm **(correct)** 3. 3:30pm **(mis)**

Figure 3 Examples for Critical Questions and Stimuli

2.2 Results

The mean of the JOLs was taken for each condition; original encoding (Fluent vs. Disfluent), misinformation phase (Fluent vs. Disfluent presentation on original

encoding). Moreover, accuracy rates per condition (Fluent vs. Disfluent) were counted. These data organizations were done via R (R Core Team, 2013) and Excel. The descriptive and data analysis of Experiment 1 was conducted on Jamovi (The Jamovi project, 2019). Before analyzing the data, the proportion of correct responses, false alarms, and the amount of false memory were calculated (*see formulas in APPENDIX C*). The difference between conditions for accuracy rates, false alarm rates, and the amount of false memory (susceptibility to misinformation) was compared across encoding conditions.

2.2.1 Memory Accuracy

The average accuracy of Experiment 1 including critical and control questions was .55. A paired samples t-test revealed that there was a significant effect for fluency, $t(55) = 4.66$, $p < .001$, $d = .62$. Fluent items ($M = .58$, $SD = .09$) produced higher memory performance than disfluent ones ($M = .52$, $SD = .1$). The accuracy comparison was also done separately for critical and control items. A paired samples t-test for accuracy percentage of critical items yielded that accuracy rates for critical items were higher in the fluent condition ($M = .56$, $SD = .12$) than disfluent condition ($M = .50$, $SD = .11$), $t(55) = 4.31$, $p < .001$, $d = .58$. Similarly, a paired samples t-test for accuracy percentage of control items showed that accuracy rates for control items were also higher in the fluent condition ($M = .30$, $SD = .04$) than disfluent condition ($M = 27$, $SD = .06$), $t(55) = 2.64$, $p < .001$, $d = .35$. However, the Cohen's d was lower for the control questions compared to critical questions. This difference might be related to the fact that the control items were presented two times whereas the critical items were presented only once due to changed version in misinformation phase.

To further investigate the memory performance, the corrected hit rates were calculated by subtracting false alarm rates (foil + misinformation) from hit rates for critical details. The comparison of fluent and disfluent conditions for corrected hit rates showed that participants produced higher hit-rates for the fluent items (M = .12, SD = .24) compared to disfluent items (M = -.006, SD = .21), $t(55) = 4.32$, $p < .001$, d = .58. Moreover, a one-sample t-test was conducted to see whether the corrected hit-rates were different from zero. A one-sample t-test revealed that corrected hit-rates in fluent condition were significantly different from 0, t(55) = 3.96, p < .001, d = .53 whereas corrected hit-rates in disfluent condition were not different from 0, indicating that participants did not respond the questions based on their recall; instead they probably guessed the answers depending on their intuition.

Besides the corrected hit rates, d' analyses were conducted to examine participants' sensitivity to differentiate the choices. d-prime (d') was calculated by subtracting the z transform of false alarms (i.e., total foil and false responses) from the hit rates (i.e., percentage of accurate responses). d' primes were calculated by using total false alarm rates (i.e., any wrong response other than correct ones) and by using false memories (i.e., the choice presented in misinformation phase) and foil responses (i.e., unrelated choice) separately as false alarms. d' prime analysis with total false alarm rates revealed a significant difference between the conditions, indicating that participants' d primes were higher in fluent condition (d' = .33) compared to disfluent condition (d' = -.01), t(55) = 4.40, p < .001, d = .59. The higher d' rates meant that participants could differentiate the correct responses more in fluent condition compared to disfluent ones. This result was in line with comparing fluent and disfluent conditions in terms of memory accuracy rates. 2 [fluency (fluent vs. disfluent)] x 2 [types of false alarms (false vs. foil)] repeated-measures ANOVA

revealed that two significant main effects of fluency and types of false alarms, showing that d' rates were higher for fluent items and d' rates were higher when only foil responses were used as false alarms. The first main effect was the replication of the t-test above, showing that participants' sensitivities were higher for fluent items than disfluent items, F(1, 55) = 18.58, p < .001, η_p^2 = .25. The other main effect showed that participants performed better at differentiating correct responses from foil items (d'_{fluent} = .51, $d'_{disfluent}$ = .72) compared to false items (i.e., items presented in misinformation phase, d'_{fluent} = -.38, $d'_{disfluent}$ = -.52), F(1, 55) = 426.68, p < .001, η_p^2 = .89, implying that misinformation manipulation might increase the confusion and possibly produced false memories. The interaction effect was not significant.

2.2.2 False Memory

As a reminder, participants chose one of the choices among three choices: a correct response, a misinformation response, or a foil response. The amount of susceptibility to misinformation was calculated by dividing the amount of false memory (i.e., number of questions in which participants chose the item presented in misinformation phase) into the total false alarm (sum of the misinformation responses and foil responses). The number of foil responses was also calculated by dividing the number of foil responses into the total false alarm. After calculating each participant's false memory and foil rates in fluent and disfluent conditions separately, 2 [fluency (fluent vs. disfluent) x 2 types of response (false vs. foil)] within-subjects design ANOVA was conducted. The significant main effect of types of response revealed that the misinformation phase created false memories, $F(1, 55) = 492.97$, $p < .001$, $\eta_p^2 = .9$. In other words, participants were more likely to choose misinformation choices ($M_{disfluent} = .69$, $M_{fluent} = .71$) compared to foil choices ($M_{disfluent} = .31$, $M_{fluent} = .30$). There was no significant main effect of fluency F(1,

55) = 3.11, p = .08, which revealed that the fluency did not change the percentage of foil and misinformation responses. Moreover, there was no interaction effect. The fluency and types of responses did not affect each other.

There were some critical findings in this study. On the one hand, there was a significant difference in the accuracy percentages between fluent and disfluent conditions. On the other hand, there was no significant difference in the false memory percentages between the conditions. Even if most of the time studies did not found a change in memory performance due to fluency manipulation (see examples in Besken & Mulligan, 2014; Rhodes & Castel, 2008, 2009), the difference in the accuracy percentages was still the replication of the findings in the literature (see examples in Yue, Castel, and Bjork, 2013; Undorf & Zimdahl, 2019). However, to-date there were not any studies that investigated the effect of fluency on the misinformation paradigm. Therefore, it was crucial to discuss the possible reasons why there was no effect in the current experiment. One of the reasons might be the high percentage rate of false memory in total ($M_{disfluent}$ = 69, M_{fluent} = 71), which might make the difference between the groups invisible, if any. The high rate of false memory might be due to the long duration of the experiment. Moreover, during Experiment 1, participants were asked to give item-by-item JOLs, which might affect memory performance in general. These reasons are discussed in a detailed way (see in 2.3. Discussion).

2.2.3 JOLs

Before analyzing, the JOLs which were not appropriately given (e.g., not in the range of 0-100, not in the number format) were eliminated from the data. Based on the rules, the .002 percentage of JOLs (44 out of 22400) in the encoding phase and .003

percentage of JOLs (63 out of 22400) in the misinformation phase were excluded from the data. As expected, a paired samples t-test with the average JOLs revealed that people were more likely to assume they would not remember the details in disfluent images well (M = 60.2, SD = 2.35) compared to fluent images (M = 63.3, SD = 12.22), $t(55) = 3.30$, $p = .002$, $d = .44$. So, perceptual disfluency made people think they would not remember the items well when they experienced difficulty in perceiving the items. This finding replicated the findings for perceptual fluency in the literature (Besken & Mulligan, 2013, 2014; Mueller et al., 2014; Reber et al., 1998; Susser et al., 2013). However, as we mentioned above, most of the time, there might be metamemory illusion due to perceptual fluency manipulation. However, participants' JOLs accurately predicted real memory performance in Experiment 1. In other words, participants did not only give lower JOLs scores to the disfluent items compared to fluent items, but they could also recall the disfluent items less accurately than the fluent items in the recognition memory task.

Table 1 JOLs of Encoding Phase in Experiment 1

(DF: Disfluent, F: Fluent, Total refers to the all details that were questioned, Critical refers to the details which changed during misinformation phase, Control refers to the details which did not change during misinformation phase)

Conditions	Mean	SD
Total F	63.3	16.6
Total DF	60.2	17.6
Critical F	66.4	15.6
Critical DF	63.0	16.6
Control F	64.1	16.2
Control DF	61.5	17.7

Table 2 JOLs of Misinformation Phase in Experiment 1

(DF: Disfluent presentation in original encoding phase, F: Fluent presentation in original encoding phase, Total refers to the all details that were questioned, Critical refers to the details which changed during misinformation phase, Control refers to the details which did not change during misinformation phase)

Conditions	Mean	SD
Total F	66.5	18.4
Total DF	64.6	18.9
Critical F	67.4	18.1
Critical DF	65.8	18.6
Control F	68.1	18.3
Control DF	65.8	19.0

To explore further, we analyzed the JOLs, which were given in the misinformation encoding phase, too. The JOLs in the misinformation encoding phase were divided into two groups as JOLs for fluently presented items in the original encoding phase and JOLs for disfluently presented ones in the original encoding phase. As a reminder, all the materials were presented fluently in the misinformation encoding phase. So, while participants were giving the JOLs, they saw the images in fluent conditions in which they were presented either in fluent or disfluent conditions beforehand. Interestingly, a paired sample t-test revealed that there was a significant difference between the fluently presented items and disfluently presented items, meaning that participants gave higher JOLs for the fluently presented items (M = 66.5, SD = 2.46) than the disfluently presented items (M = 64.6, SD = 2.52), $t(55)$ = 2.18 , p = .03, d = .29. The results might be important to show that people's first judgment for material could affect later judgments for the same material.

Since the materials were presented fluently, the overall JOLs' means in the misinformation phase were higher than the encoding phase. However, it might be critical to understanding whether the increase in JOLs changed across groups (fluently presented vs. disfluently presented in first encoding) in order to reach further interpretation. A 2 [first presentation fluency (fluent vs. disfluent)] x 2 [presentation time (first vs. second)] repeated-measures ANOVA was conducted. There was a main effect of first presentation fluency; fluent items were given higher JOLs, $F(1, 55) = 9.24$, $p = .004$, $\eta_p^2 = .14$. There was also the main effect of presentation time; the items were given higher JOLs in the second presentations $F(1, 55) = 7.53$, $p = .008$, $\eta_p^2 = .12$. There was no significant interaction effect between the variables. The insignificant interaction effect showed that the first presentation fluency did not influence the increase in the second presentation.

Another important aspect of the current study was that there were changed details (critical) and unchanged details (control). People made JOLs for both the unchanged and changed items during the misinformation phase. For further exploration, it was important to analyze whether people's JOLs differed across critical and control items from first to second presentation time based on the presentation fluency. To examine the question, a 2 [first presentation fluency (fluent vs. disfluent)] x 2 [presentation time (first vs. second)] x 2 [item types (critical vs. control)] repeated-measures ANOVA was conducted, which replicated the results of the 2x2 ANOVA above. The results revealed a significant main effect for the presentation time and the first presentation fluency. Participants made significantly different JOLs from first to second presentation $F(1, 55) = 10.1$, $p = .002$, $\eta_p^2 = .15$. and for fluent and disfluent. Their JOLs also increased from the first presentation to the second presentation. There were no significant differences between the critical and control items in terms

of JOLs. The results of the 2x2x2 ANOVA did not yield any significant interaction effect among the variables.

2.3 Discussion

Depending on these results of JOLs, it might be concluded that disfluent presentation produced a decrease in JOLs, repeated exposure caused an increase from the first presentation to the second presentation and participants' JOLs did not differ according to the item types in general. The result of fluency was the replication of the general findings in the literature (see in Besken & Mulligan, 2014; Rhodes & Castel, 2008) and affirmed the hypotheses related to JOLs. Moreover, the increase in the second presentation might be explained by the mere-exposure effect, which made the items more fluent and produced higher JOLs. The non-significant result of critical and control items comparison showed that the details' changes did not influence people's JOLs in the second presentation. Participants might not be aware enough to change their judgments for the critical items.

In addition to these critical findings, another important aspect of the results was the difference between the effect of fluency on accurate and false memories. Whereas the performance was decreased through the disfluent presentation for accurate memories, the number of false memories, in other words, susceptibility to misinformation, did not change across fluent and disfluent presentations. Moreover, d' analyses also revealed that participants' accuracy rates were higher for fluent conditions compared to disfluent ones, and participants were more likely to differentiate the correct choices from foil choices compared to false choices, which supporting the memory accuracy findings. The results supported the hypotheses

related to actual memory performance, but not the false memory. This difference might have originated from providing JOLs throughout the experiment. In the literature, the effect of providing JOLs while learning the material on the subsequent memory performance for the material was named as *JOL reactivity*. There are a wide variety of results related to the JOL reactivity phenomena. Previous findings supported that providing JOLs produce higher memory performance in the final cued-recall test (Myers, Rhodes, & Hausman, 2020; Rhodes & Tauber, 2011; Soderstrom, Clark, Halamish, & Bjork, 2015). Witherby and Tauber (2017) maintained that improvement of the making JOLs during encoding might be observed in both short-term memory test (2 min after the study) and long-term learning test (3 days after the study). However, the reactivity effect might be observed depending on the methodological structure of the experiments. For instance, Janes, Rivers, and Dunlosky (2018) used related and unrelated pairs as the stimuli and found that the making JOLs increased the relatedness effect more in the mixed list compared to the pure list.

In contrast to these findings, there are also some studies in which the making JOLs did not produce any further improvement in the subsequent memory test. Soderstrom et al. (2015) observed the JOLs improved the cued-recall performance just for the related pairs, but not for the unrelated ones. Furthermore, other studies did not find any change in memory performance due to JOLs (e.g., Begg, Martin, & Needham, 1992; Tauber & Rhodes, 2012). Unlike Rhodes and Tauber (2011), Double, Birney, and Walker (2018), in their meta-analysis, found that there was no effect of making JOLs during encoding on the future memory performance regardless of the methodological settings. They found the reactivity effect when they created sub-groups (e.g., memory performance for related pairs). These findings suggested that it

was essential to control the effect of JOLs on memory performance to eliminate the possible unknown effect of JOLs, if any. Therefore, I decided to conduct Experiment 2 with some changes.

CHAPTER 3

EXPERIMENT 2

Since the results of Experiment 1 might have been affected from JOL reactivity and the long duration of the experiment, we decided to eliminate JOLs and shorten the duration. This, in turn, might also eliminate the reactivity effect for JOLs, if any, since the mere act of making JOLs might have an effect on the memory performance in terms of the accuracy and false memory rates (Janes et al., 2018). The method was identical to Experiment 1. The only difference was that metacognitive judgments were excluded from this study for both the encoding and misinformation phase, which decreased the duration for half an hour. The main hypothesis was people would have less accurate and more false memories for disfluent items compared to fluent items.

3.1 Method

3.1.1 Participants

Fifty-six students *(34 females,* M_{age} *= 19.65,* SD *= 1.21)* from I. D. Bilkent University participated in exchange for course credits. Participants were all native

Turkish speakers and had normal or corrected to normal vision.

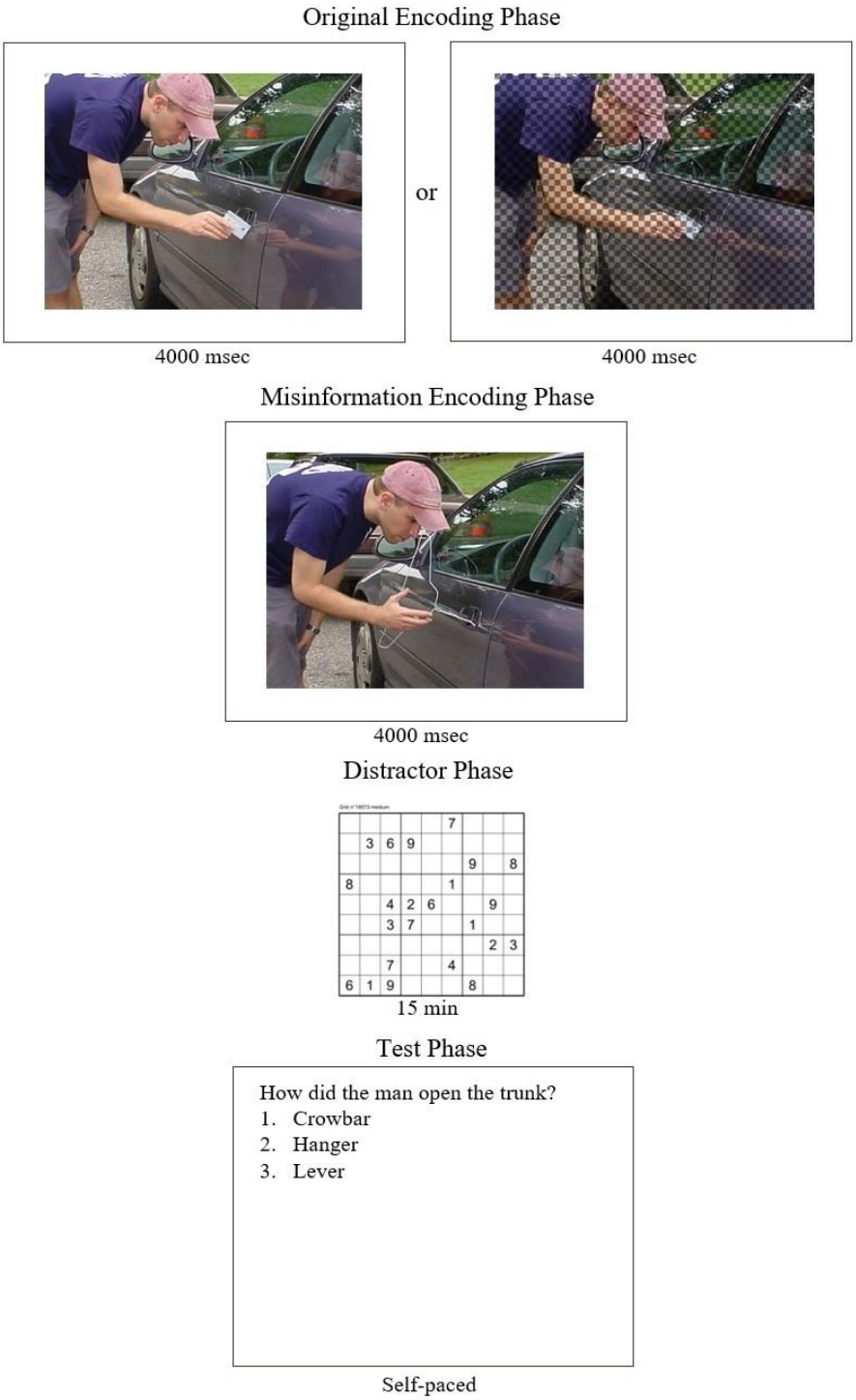


Figure 4 Four Main Stages of Experiment 2: Original Encoding Phase, Misinformation Encoding Phase, Distractor Phase, Test Phase.

3.1.2 Procedure

The materials were identical to Experiment 1. The procedural details were similar to Experiment 1. In this experiment, the same materials were presented in the same way. Different from Experiment 1, the JOL procedure was eliminated from the study, which shortened the duration of the experiment from two hours to one and a half hours and leads to a more continuous presentation of the story (see in Figure 2).

3.2 Results

The descriptive and data analysis of Experiment 2 was also conducted on Jamovi. Before analyzing the data proportion of correct responses, false alarms, and the amount of false memory were calculated via R. The difference between conditions for accuracy rates, false alarm rates, and the amount of false memory was compared.

3.2.1 Accuracy

The total average of accuracy percentage in Experiment 2 was .53. A paired samples t-test revealed that the effect of fluency on accuracy replicated the results of Experiment 1. The fluent items (M = .56, SD = .09) produced higher memory performance than disfluent ones (M = .49, SD = .07), $t(55) = 6.61$, $p < .001$, $d = .8$. The effect size of the fluency effect on accuracy was larger than the Experiment 1. The JOLs might be the reason for this difference.

As in Experiment 1, the accuracy comparison was also conducted separately for critical and control items. The results of two separate paired samples t-tests for critical and control items revealed a significant difference between fluent and disfluent conditions. For the critical items, critical items in fluent condition (M = .56,

SD = .1) produced higher memory performance compared to the critical items in disfluent condition (*M* = .47, *SD* = .09), *t*(55) = 6.12 , *p* < .001, *d* = .8. Likewise, the control items in the fluent condition (*M* = .29, *SD* = .05) produced higher memory performance than control items in the disfluent condition (*M* = .26, *SD* = .04), *t*(55) = 2.29 , *p* < .001, *d* = .3. As expected, in Experiment 2, the effect size of critical items difference was also larger than the effect size of control items. The reason for the small effect size might be the repetitive presentation of control items while the critical items were presented only once.

As in the analysis of Experiment 1, the corrected hit rates were used for the further investigation of the memory performance, specifically for critical details. The corrected hit rates comparison for the fluent and disfluent conditions manifested higher hit-rates for the fluent items (*M* = .12, *SD* = .23) compared to disfluent items (*M* = -.006, *SD* = .2), *t*(55) = 6.16 , *p* < .001, *d* = .82. In a similar vein, a one-sample t-test was conducted to compare the corrected hit-rates with zero. A one-sample t-test showed that the result of Experiment 2 was similar to Experiment 1 for fluent condition, in which corrected hit-rates were significantly different from zero, *t*(55) = 3.85, *p* < .001, *d* = .51. In contrary to Experiment 1, corrected hit-rates in disfluent conditions were also significantly different from zero in Experiment 2, *t*(55) = -2.40, *p* = .02, *d* = .32.

In addition to these analyses, we conducted d' prime analyses in the same way with Experiment 1. A paired samples t-test revealed that participants' d primes were higher in fluent condition (d'_{fluent} = .32) compared to disfluent condition (d'_{fluent} = -.16), t(55) = 6.11, p < .001, d = .59. This result implied that participants could differentiate the correct responses more in fluent condition compared to disfluent ones. Moreover, 2 [fluency (fluent vs. disfluent)] x 2 [types of false alarms (false vs.

foil)] repeated-measures ANOVA replicated the results of Experiment 1. Participants' d primes (i.e., sensitivity while differentiating the correct and wrong choices) were higher for fluent items than disfluent items, F(1, 55) = 36.7, p < .001, η_p^2 = .4. Also, d primes were higher for foil items (d'_{fluent} = .63, $d'_{disfluent}$ = .46) than false items (i.e., items presented in misinformation phase, d'_{fluent} = -.30, $d'_{disfluent}$ = -.61), F(1, 55) = 275.21, p < .001, η_p^2 = .83, implying that misinformation manipulation might increase the confusion and possibly produced false memories. The interaction effect was not significant as in Experiment 1.

3.2.2 False Memory

2 [fluency (fluent vs. disfluent) x 2 types of response (false vs. foil)] within-subjects design ANOVA was conducted after calculating each participant's false memory and foil rates in fluent and disfluent conditions separately. There was a main effect of response type which revealed that the misinformation manipulation created significant amount of false memory $F(1, 55) = 350.13$, $p < .001$, $\eta_p^2 = .86$. The misinformation choices ($M_{disfluent}$ = .69, M_{fluent} = .67) were chosen more than the foil choices ($M_{disfluent}$ = .33, M_{fluent} = .31). The main effect of fluency was not significant, meaning that the fluency did not influence the tendency to remember foil and false choices as correct answers. Since the interaction effect was not also significant, it could be concluded that fluency of the information did not change the amount of false and foil responses.

Even if the results of the Experiment 2 replicated the results of Experiment 1, it was critical to see whether JOLs had a further effect on the results. Since it was not planned to analyze the effect of JOLs across the experiments, we did not control for the time spent for giving JOLs during encoding phases. However, to reach further

understanding, we compared the two experiments with each other. Interestingly, the means for the accuracy ($M_{disfluent1}$ = 52, $M_{disfluent2}$ = 49, $M_{fluent1}$ = 58, & $M_{fluent2}$ = 56.4), false choices ($M_{disfluent1}$ = 69.3, $M_{disfluent2}$ = 69.4, $M_{fluent1}$ = 71, & $M_{fluent2}$ = 67) and foil choices ($M_{disfluent1}$ = 30.7, $M_{disfluent2}$ = 30.6, $M_{fluent1}$ = 29.5, & $M_{fluent2}$ = 32.6) did not differ across experiments.

3.3 Discussion

Experiment 2 was conducted without JOLs to control the possible effect of making JOLs. In this experiment, misinformation manipulation significantly created false memories in the subsequent memory test. Moreover, in parallel with the results of Experiment 1, the results of Experiment 2 showed that disfluent presentation decreased the accurate memories but did not influence the false memories. Results suggested that a lack of making JOLs did not change the direction of the results. For the further comparison of false and accurate memories in Experiment 1 with JOLs and Experiment 2 without JOLs, an independent sample t-test with the percentages of accurate ($M_{experiment1fluent}$ = 0.58, $M_{experiment2fluent}$ = .56, $M_{experiment1disfluent}$ = 0.51 $M_{experiment2disfluent}$ = .48) and false memories ($M_{experiment1fluent}$ = 0.71, $M_{experiment2fluent}$ = .67, $M_{experiment1disfluent}$ = .69 $M_{experiment2disfluent}$ = .69) revealed no significant different between the experiments. Moreover, in Experiment 1, there might be a fatigue effect since the experiment took two hours. In Experiment 2, even though the duration of the experiment decreased, the means of accuracy rates in fluent and disfluent conditions also decreased.

In the comparison of corrected hit-rates with zero, we found a negative correlation, which suggested that people responded to the questions significantly below the chance level, showing that disfluent presentation produced lower memory than

chance level. However, in the previous experiments, there were 400 stimuli for the initial encoding phase and 400 stimuli for the misinformation encoding phase, which might create confusion due to lots of unrelated details that were not used as critical or control. The number of unrelated items might decrease the participants' attention for the critical and control details, which were asked in the final memory tests. Furthermore, d' analyses revealed that d primes were higher for fluent condition compared to disfluent condition, indicating that participants' sensitivities (i.e., how accurate they differentiate the correct choices from wrong choices) were higher in fluent condition than disfluent condition. Due to the high amount of noise, participants might confuse and not learn the critical details well. The decrease in accuracy in disfluent conditions might also be explained through the effect of noise. I expected to observe an increase in false memories due to noise; however, the effect of noise might increase the overall false memories, which might eliminate the difference between the fluent and disfluent conditions. To investigate further, I decided to conduct Experiment 3 with a smaller number of stimuli; therefore, I decreased the number of vignettes and the number of images in each vignette, which would decrease the number of unrelated details and the duration.

CHAPTER 4

EXPERIMENT 3

The results of Experiment 2 showed that both the low accuracy level and high false memory rate stayed the same with Experiment 1, even if we decreased the time from two hours to one and a half hours and removed the JOL task. It meant that there was no JOLs effect on memory performance. In other words, we did not observe the reactivity effect. However, the number of false memories were still high for the fluent and disfluent conditions. Therefore, it was important to investigate whether the higher false memory rates prevented the acquisition of the fluency effect on false memory through the misinformation paradigm. Therefore, in Experiment 3, two modifications to the experimental procedure were made.

In the previous experiments, there were many details in the images which were unrelated to critical details. The twelve images were used as critical images, and six images were used as control images, and the remaining twenty-two images were just unrelated images to the final memory test. These unrelated images might produce some confusion. They might increase memory confusion through proactive and retroactive interference since they increased the noise in the stimuli set, which might

result in higher false alarm rates. Therefore, the first change was decreasing the number of unrelated pictures. While decreasing the number of unrelated pictures, the logical flows of the stories were controlled by asking pilot participants. Moreover, in both studies, participants were forced to respond during the 3-choice recognition-test, which might produce guess responses. Since the distribution of the amount of guess response into the fluent and disfluent groups was unknown, it was also critical to control the effect of guess responses. Thus, in the current study, the small box was presented after each image during the recognition phase, and participants were asked to indicate that they guessed the response by filling in the box after answering the question.

4.1 Method

4.1.1 Participants

Seventeen students (9 females, M_{age} = 20.88, SD = 1.97) from I. D. Bilkent University participated in exchange for course credits. Participants were all native Turkish speakers and had normal or corrected to normal vision.

4.1.2 Materials

In this experiment, the materials were presented in the same manner. Different from Experiment 2, we decreased both the number of stories and the number of images in each story. The number of stories was decreased from 8 to 4. The number of images in each story was also decreased from 50 to 30. In this case, participants were presented with 240 images (in total during original and misinformation phases) instead of 800 images. The order of conditions was partially counter-balanced,

depending on these changes (*see in APPENDIX D – b*). The same rules with Experiments 1 and 2 were used. As in Experiment 2, we did not collect JOLs to prevent an increase in the duration. Due to the decreased number of images, Experiment 3 took 40-45 minutes only.

Four stories for Experiment 3 were chosen according to the flow of stories and vividness of the details. In other words, we presented stories to the people of the independent group and asked which details they remembered more. Moreover, we paid attention to hold the flow of stories logically while choosing 30 images for each story. At the end of the process, each story had 30 images with 12 critical, six control, and 12 filler images.

4.1.3 Procedure

The procedural details were similar to previous experiments. Different from the previous experiments, to control guessed responses, we asked participants to report the questions that they thought they did not see the detail and gave their responses based on intuition during the recognition phase. They reported the questions with the guessed responses by filling a box that was presented after the participant gave a response to the recognition test question. For instance, a participant did not see which object a man in the story was holding in his hand. First, we asked her/him to answer the question, depending on their intuition. Then, they wrote the “x” letter to the box to mark the box when they were confident that they could not see the detail. During this phase, participants were informed about not marking the box for the questions that they could not remember or be sure. Exact instruction for guess response which was used during the experiment is below:

"Please, fill the box which will appear just after you respond the question if only you think you could not see the detail that was asked for and you were guessing the answer intuitively. Please, do not fill the box when you think you did not recall accurately, or/and you were not sure about your answer."

4.2 Results

The analyses were made both with and without box-control. The without box-control analysis refers to the analyses including all the responses participants gave; whereas, with box-control refers to the analyses in which the guess responses (i.e., the responses participants wrote x to the box) were subtracted from the data. Before the subtraction, there were 36 responses in each condition per participant; the average total number of questions in fluent and disfluent conditions fell into 29.2 (21.1 critical and 8.1 control) and 28.6 (20.7 critical and 7.8 control), respectively.

4.2.1 Without Box Control

Since Experiment 3 consisted of 4 stories with 30 images, the results without box control were also critical to examine whether decreasing the number of stimuli and the noise due to unrelated images with critical details influenced the results or not. As a reminder in this analysis, all responses participants gave, including guess responses (i.e., participants filled the box) were used; therefore, there were 72 responses (36 fluent and 36 disfluent) in total. The without box-control analysis provided a chance to compare the results with the results of Experiments 2 since the only difference from it was the number of stories and images.

4.2.1.1 Accuracy

The total average of accuracy percentage in Experiment 3 without box-control was .59. As expected, the overall accuracy rate was higher than the previous experiments since the number of filler images in Experiment 3 was lower than the previous ones. The results of a paired samples t-test demonstrated that participants were more likely to recall the fluent items (M = .66, SD = .12) accurately compared to disfluent ones (M = .52, SD = .09), $t(16) = 4.13$, $p < .001$, $d = 1$. Even if the small number of participants, Cohen's d was higher than the previous experiments, certifying that the difference between fluent and disfluent conditions in terms of the accuracy rate was observed better with a lower number of stories and lower number of images in each story.

The separate analysis for accuracy rates of critical and control items produced results similar to previous experiments. Critical items in the fluent condition (M = .65, SD = .14) were more likely to be remembered accurately compared to the critical items in the disfluent condition (M = .53, SD = .14), $t(16) = 2.85$, $p = .012$, $d = 0.69$. Moreover, control items in fluent condition (M = .67, SD = .15) were more likely to be remembered accurately compared to the control items in the disfluent condition (M = .50, SD = .11), $t(16) = 3.68$, $p = .002$, $d = 0.89$. As expected, the effect size of the analysis for critical items were higher than the effect of the analysis for control items that were repeatedly presented. Accuracy rates comparison for fluent versus disfluent presentations replicated the findings of previous experiments.

As in the previous analysis, the corrected hit rates were used for further investigation of the accuracy rates of critical details. A t-test for the comparison between fluent and disfluent conditions showed higher hit-rates for the fluent items (M = .30, SD =

.29) compared to disfluent items (M = .05, SD = .28), $t(16) = 2.85$, $p < .012$, $d = .69$. Since one of the aims of Experiment 3 was to increase the accuracy rates, a one-sample t-test for comparison of corrected hit-rates with zero was critical to examine so that we can see whether the corrected hit-rates were different from zero (i.e., chance level). A one-sample t-test revealed that the corrected hit-rates in fluent condition was significantly different from 0, $t(16) = 4.42$, $p < .001$, $d = 1.071$. In contrast the corrected hit-rates in disfluent conditions were not significantly different than 0, $t(16) = 0.78$, $p = .44$, $d = .19$ in Experiment 3 without box control, showing that participants' responses were still based on intuition mostly.

Similar to previous experiments, d' prime analyses revealed that in fluent condition (d' = .85) d primes were higher compared to disfluent condition (d' = .14), $t(16) = 2.94$, $p = .01$, $d = .71$, showing that participants' were more sensitive while differentiating correct and wrong choices in fluent condition than disfluent. Moreover, 2 [fluency (fluent vs. disfluent)] x 2 [types of false alarms (false vs. foil)] repeated-measures ANOVA showed participants' d primes were higher for fluent items than disfluent items, $F(1, 16) = 8.61$, $p < .01$, $\eta_p^2 = .35$. Moreover, participants' d primes were higher for foil items ($d'_{fluent} = 1.37$, $d'_{disfluent} = .99$) than false items ($d'_{fluent} = -.51$, $d'_{disfluent} = -.85$), $F(1, 16) = 58.9$, $p < .001$, $\eta_p^2 = .79$. The other main effect indicated that misinformation manipulation might increase the probability to confuse false choice (i.e., the item presented in misinformation phase) with correct choice, which refers to the susceptibility to misinformation. The interaction effect was not significant, as found in previous experiments.

4.2.1.2 False Memory

The rates of false and foil responses of participants were calculated separately for fluent and disfluent conditions. 2 [fluency (fluent vs. disfluent) x 2 types of response

(false vs. foil)] within-subjects design ANOVA revealed that participants choose the false responses ($M_{disfluent}$ = .80, M_{fluent} = .82) significantly more than the foil response ($M_{disfluent}$ = .18, M_{fluent} = .20) in the final recognition test, $F(1, 16) = 88.4$, $p < .001$, $\eta_p^2 = .85$. The significant difference between the false and foil response showed misinformation manipulation created false memories, and participants assumed that they saw the changed details in the first presentation even though they were presented with misinformation during the misinformation phase.

One of the aims of the current project was to examine whether disfluent presentation increased the false memory rates, or not. The results showed that there was no effect of fluency on the rate of false memories. Moreover, there was no interaction between the fluency and the probability of choosing foil or false items. It might be concluded that even if the duration of Experiment 3 was shortened, and the number of noises decreased, the disfluent presentation did not change the number of false memories. Besides, the shorter duration and the decrease in the noises increased the accuracy rates.

4.2.2 With Box Control

In the analysis with box control, the responses in which participants reported they guessed the response by filling the box were excluded from the data. 18% of data in fluent condition and 20% of data in disfluent conditions were excluded. After excluding the guess responses, the calculations and analyses were done with the rest of the questions in fluent and disfluent conditions, and the average number of questions in fluent and disfluent conditions were 29.2 (21.1 critical and 8.1 control), and 28.6 (20.7 critical and 7.8 control), respectively.

4.2.2.1 Accuracy

The total average of accuracy percentage in Experiment 3 with box-control was .65. The overall accuracy rate was higher than all previous findings, including without box-control analysis of Experiment 3. In this analysis, we controlled both for unrelated images and the wild guess responses. A paired samples t-test showed that the effect of fluency on accuracy replicated the results of previous experiments and with box control analyses of Experiment 3. The fluent items (M = .73, SD = .14) produced higher memory performance than disfluent ones (M = .56, SD = .14), $t(16) = 4.22$, $p < .001$, $d = 1.02$. The effect size of the fluency effect on accuracy was larger than all previous results despite the smaller number of participants.

Accuracy rates of fluent and disfluent conditions were also compared separately for critical and control items. The analyses showed that ,as in the previous results, fluently presented critical items (M = .69, SD = .16) were remembered more than the disfluently presented critical items (M = .55, SD = .17), $t(16) = 3.17$, $p = .006$, $d = 0.76$. The same analysis for control items produced similar results, showing that fluently presented control items (M = .81, SD = .17) were more likely to be recalled accurately than the disfluently presented control items (M = .60, SD = .16), $t(16) = 3.97$, $p = .001$, $d = 0.96$. Interestingly, the effect size of the analysis for control items was larger than the effect of the analysis for critical items. The change in the direction of the effect size compared to the previous studies might be due to the control of guess responses. A t-test comparison across fluent and disfluent conditions with corrected hit rates of critical responses showed that higher hit-rates for the fluent items (M = .38, SD = .32) compared to disfluent items (M = .09, SD = .34), $t(16) = 3.17$, $p = .006$, $d = .76$. In Experiment 3 with box control, accuracy rates increased compared to all other conditions and experiments; therefore, a one-sample

t-test for comparison of corrected hit-rates was critical to examine whether the corrected hit-rates were different from zero. The result of a one-sample t-test showed that the corrected hit-rates in fluent condition was significantly different from zero, $t(55) = 4.89$, $p < .001$, $d = 1.19$; however, the corrected hit-rates in disfluent conditions were not significantly different in Experiment 3 with box control. The result pointed out that participants were still answering intuitively even if their overall accuracy rates increased. As in all previous analyses, in Experiment 3 with box-control analysis, d' primes did not differ across fluent and disfluent conditions for total false alarm rates ($p > .05$), false responses ($p > .05$), and foil responses ($p > .05$).

In Experiment 3 with box-control analysis, it was interesting to conduct d' prime analyses since the overall accuracy rates increased both for fluent and disfluent items. The results replicated the previous experiments and the condition. A paired sample t-test manifested that d primes were higher in fluent condition ($d' = 1.12$) compared to disfluent condition ($d' = .27$), $t(16) = 3.57$, $p = .002$, $d = .84$. Participants' differentiations were more accurate in fluent condition compared to disfluent. Furthermore, 2 [fluency (fluent vs. disfluent)] x 2 [types of false alarms (false vs. foil)] repeated-measures ANOVA revealed that d primes were higher for fluent items than disfluent items, $F(1, 16) = 11.4$, $p = .004$, $\eta_p^2 = .41$. Participants' d primes were also higher for foil items ($d'_{fluent} = 1.74$, $d'_{disfluent} = 1.24$) than false items ($d'_{fluent} = -.62$, $d'_{disfluent} = -.98$), $F(1, 16) = 41.7$, $p < .001$, $\eta_p^2 = .72$. The results indicated that accuracy was higher in fluent condition, and participants differentiated correct response more easily in fluent condition than disfluent one. The misinformation manipulation might increase the probability to confuse false choice

(i.e., the item presented in misinformation phase) with the correct choice. The interaction effect was not significant.

4.2.2.2 False Memory

Each participants' rates of false and foil responses were calculated for fluent and disfluent conditions, separately. 2 [fluency (fluent vs. disfluent) x 2 types of response (false vs. foil)] within-subjects design ANOVA revealed that the false responses ($M_{disfluent}$ = .80, M_{fluent} = .82) were chosen significantly higher than the foil response ($M_{disfluent}$ = .18, M_{fluent} = .20), $F(1, 16) = 88.4$, $p < .001$, $\eta_p^2 = .85$, meaning that false memory manipulation worked properly. However, there was no main effect of fluency manipulation on the probability to choose foil or false response. There was also no significant interaction effect between fluency and types of response, showing that fluent or disfluent presentation did not change the number of false memories.

Even if the possible effect of guess responses were controlled in Experiment 3 (see in with box control part), all the results were in line with the previous experiments. Since the results were the same, it might be interpreted that there is no direct effect of guess responses on the results of previous experiments. To find out proof from the data for this interpretation, and 2 (box control) x 2 (fluency) x 2 (types of response) ANOVA was conducted with the expectation of no significant interaction effects between any combination of variables. The effect of box control and the interaction effects were not significant, revealing that the guess responses did not influence the results (see F values and p values in Table 3).

Table 3 Results of 2 (box control) x 2 (fluency) x 2 (types of response) ANOVA

	F	*p*	η^2_p
Box Control	-1.54e−13	1.00	-0.00
Fluency	9.11e-14	1.00	0.00
Types of Response	109.7524	< .001	0.87
Box Control ∗ Fluency	-1.11e−14	1.00	-0.00
Box Control ∗ Types of Response	3.6678	0.07	0.18
Fluency ∗ Types of Response	0.2341	0.64	0.01
Box Control ∗ Fluency ∗ Types of Response	0.0199	0.89	0.001

Table 4 Means of Accurate and False Memories in Experiment 1, Experiment 2, and Experiment 3

(The numbers in parentheses refers to the standard deviations)

a. Accuracy Percentage and Corrected Hit-Rates of 3 Experiments

	Accuracy Percentage		Corrected Hit-Rates	
Conditions	Fluent	Disfluent	Fluent	Disfluent
Experiment 1	.58 (.09)	.52 (.1)	.12 (.24)	-.006 (.21)
Experiment 2	.56 (.09)	.49 (.07)	.12 (.23)	-.006 (.2)
Experiment 3 - with box	.72 (.14)	.56 (.13)	.38 (.32)	.09 (.34)
Experiment 3 - without box	.66 (.1)	.52 (.09)	.30 (29)	.05 (.28)

b. False Memory Percentages of 3 Experiments

	False Memory Percentage	
Conditions	Fluent	Disfluent
Experiment 1	.69 (.08)	.70 (.08)
Experiment 2	.67 (.09)	.69 (.09)

Table 4 (cont'd)

Experiment 3 - with box	.82 (.15)	.80 (.16)
Experiment 3 - without box	.79 (.1)	.77 (.1)

4.3 Discussion

Experiment 3 was conducted to control the number of unrelated details and guess responses to examine whether these two variables had an effect on previous experiments. I was planning to collect data from 56 participants to reach .85 power depending on the power analysis; however, the data collection process was interrupted due to Covid-19. I decided to report the data from 17 participants. In the future, further research might be conducted to reach appropriate power.

Despite the small number of unrelated details and control for guess responses, the results of Experiment 3 replicated the results of Experiment 1 and 2. The results affirmed that participants recall more accurately in the fluent condition than in the disfluent condition; however, their false memories did not differ across conditions, as also found in previous experiments. The analyses with corrected-hit rates and d primes also replicated the results of previous experiments. According to these results, it can be interpreted that the encoding process was corrupted through the disfluent presentation; however, the decrease in accuracy did not increase false memories. I compared the results with box control and results without box control, which showed accuracy rates increased when the guess responses were eliminated; however, there was not any difference between the direction of the results.

CHAPTER 5

GENERAL DISCUSSION

5.1 Interpretation of Results and Theoretical Explanations

The current thesis project examined how fluent and disfluent presentation of the information influences memory predictions, actual memory performance, and false memories. There were three main questions: whether memory predictions and their accuracy would change when the information was presented in an intact or masked (1), whether actual memory performance would change depending on the fluency manipulation (2), format whether the perceptual disfluencies during the original encoding of an event may make participants more susceptible to misinformation if the misinformation is presented in a perceptually fluent manner (3). In three sets of experiments, participants were presented with stories consisting of images (eight stories with fifty images in Experiment 1 and 2; four stories with thirty images in Experiment 3). Half of the stories were shown fluently, in their intact form, while the other half were shown with the checker-board mask to make the images perceptually disfluent. Participants were asked to study these images carefully. After the encoding

phase, misinformation manipulation was used to create false memories. In the misinformation phase, participants saw all the stories that they saw in the encoding phase again in their intact form with twelve details' changes (e.g., the tool that guy used to break the door was changed from a hanger to a credit card). Participants were informed to pay attention to the stories during this phase as well because they were deceived that the true nature of the experiment was to examine the effect of repeated study sessions on predicted and actual memory. In the end, they were tested on their memories through a forced-choice recognition test. For the test session, they were informed to use only the details from the original encoding phase.

In Experiment 1, additionally, participants were asked to rate their future memory performance from a scale ranged from 0-100 (not confident I would remember quite confident I would remember, respectively). JOLs were collected to see how the metamemory processes would make a change in the relationship between fluency and false memory. In Experiment 2 and 3, participants did not make JOLs so that the effect of making JOLs, if any, can be investigated. Experiment 3 examined the possible effect of guess responses and noises due to a high number of unrelated items. Since the recognition test was forced-choice, in Experiment 1 and 2, participants could guess when they could not recall the exact information. Therefore, participants were presented with a box after each response. They were still required to respond, but they also needed to report which responses they guessed by filling the box. Moreover, since there were a high number of images (800 in total) with lots of information, there were a high number of unrelated details that were not asked in memory test and might create noises.

Results showed that participants made lower memory prediction for the disfluent items and recalled them less accurately compared to fluent ones, indicating that they

had relatively accurate memory predictions for their performance. In other words, a metacognitive illusion of perceptual fluency was not observed in the current project. Moreover, Experiment 2 and Experiment 3 replicated the findings of Experiment 1, which showed that disfluency influenced the actual and predicted memory in the same direction. Therefore, it can be concluded that making JOLs and the guess responses did not change the results. These results of JOLs were in line with the perceptual fluency hypothesis, which argues people assume they would not remember disfluently presented items compared to the fluently presented items regardless of the actual memory performance (Besken & Mulligan, 2013, 2014; Rhodes & Castel, 2008, 2009).

Interestingly, the previous findings for the effect of fluency on actual and predicted memory mostly predicted a dissociation between predicted and actual memory performance. The possible dissociation between the actual and predicted memory happened due to metamemory illusion (Rhodes & Castel, 2008) or desirable difficulty (Bjork, 1994). The former suggested that people might predict lower memory performance in disfluent conditions; however, manipulation for fluency actually did not affect the memory. These examples were mostly observed in the perceptual fluency manipulation that did not influence whether the person learns the object or not (e.g., Rhodes & Castel, 2008). Rhodes and Castel (2008) manipulated the perceptual fluency through font size, and people assumed the smaller font size decreased the probability of recalling the information compared to the larger font size. The latter argued that if the fluency manipulation increased the effort participants invested while studying, the difficulty produced higher memory performance for the hard-to-learn (disfluent) items. For instance, even if reading the information was an easier task than retrieving the information, Karpicke et al. (2009)

found that the memory performance was higher for the retrieving compared to re-reading them due to deeper (Alter et al., 2007), and more careful encoding (Song & Schwarz, 2008). There were also some examples that the effect of fluency manipulation congruently influenced the actual and predicted memory (Dunlosky & Nelson, 1992); however, they mostly found that delayed JOLs for predicting the future memory performance was more accurate compared to immediate JOLs. However, the participants made immediate item-by-item JOLs, and they accurately predicted their future performance.

There might be some possible explanations for the partially inconsistent findings with the literature. Koriat and Bjork (2005) proposed people used intrinsic and extrinsic cues to make JOLs, which influence the accuracy of JOLs. Moreover, processing and perceptual fluency were introduced as one of the immediate cues for JOLs (Jacoby & Kelley, 1987; Rhodes & Castel, 2008). Therefore, the first possible explanation for the high accuracy in making JOLs might be related to the decrease in the actual memory performance due to disfluency manipulation. In other words, the perceptual fluency cue served as a helpful cue for the reality in which participants' performances for disfluent materials were lower than those for fluent materials. Even if the pilot study was conducted to control whether participants perceived the material accurately in disfluent presentation, the disfluency manipulation might still distort the encoding process. The second explanation might be related to the accuracy increase in the process of making JOLs. In other words, people might use different cues other than perceptual fluency. They might accurately monitor the strength of the encoding process. In the literature, it was supported that even if the cues for JOLs might be misleading (Benjamin, Bjork, & Schwartz, 1998), Koriat and Shitzer-Reichert (2002) proposed that JOLs accuracy increased when people used the

internal cues as the base for them. Another explanation might be related to the effect of practicing in making JOLs. In the current study, participants gave 400 JOLs in the original encoding phase, followed by 400 more JOLs in the misinformation phase. Koriat (1997) also found that people used internal cues more when they were exposed to repeated study cycles, which might be the reason for accurate JOLs in the current study.

Moreover, further analysis of JOLs revealed some critical points. The first one was that, in the misinformation phase, participants gave higher JOLs to the items that were previously presented fluently in the encoding phase compared to the items that were presented disfluently in the encoding phase, although all items were presented fluently in the misinformation phase. In other words, while people were giving JOLs for the items, they saw the items in fluent condition; however, they made JOLs based on fluency in the initial encoding phase. The types of stimuli might explain this continuing effect of fluency on JOLs. The reason for this interesting finding might be story-like stimuli.

Furthermore, participants made higher JOLs in the misinformation phase compared to the initial encoding phase regardless of the first presentation fluency. This increase was expected for the disfluent items since they were initially presented in disfluent conditions, followed by fluent. As it was mentioned above, people's JOLs for the fluent items are higher than the disfluent ones. An interesting thing in these findings was an increase in JOLs for fluent items. This result might be explained through the mere-exposure effect, which refers that repeated exposure to a stimulus produced perceptual fluency by increasing the ease of processing (Bornstein & D'Agostino, 1994). The mere-exposure has an impact on people's judgments on the stimuli (Reber et al., 1998; Tiede, Derksen, & Leboe, 2009). More specifically, Tiede et al.

(2009) found that repeated exposure to a stimulus produced higher JOLs than a one-time presentation.

Another aim was to investigate whether the perceptual fluency of the stimuli during encoding changed the amount of false memory, or susceptibility to false information during the misinformation phase, in the subsequent memory test. I hypothesized that participants would be more inclined to integrate details from the misinformation phase to their memory if they did not clearly see the information in the original encoding phase, resulting in higher levels of false memory for disfluent than fluent stories. The basic logic behind the hypothesis was that people's beliefs about their future memory performance might change their tendency to learn the information during the second presentation and/or disfluent presentation might directly distort memories, which increases the memory errors. In the literature, it was found several times that people were more likely to give higher JOLs for the fluent items than the disfluent items (Besken & Mulligan, 2013; Rhodes & Castel, 2009). Most of the time, these predictions might produce certain behaviors to improve future recall, such as restudying the item with low JOLs (Hanczakowski et al., 2014), studying the item with low JOLs longer (Undorf & Ackerman, 2017); directing attention accordingly (Marsh et al., 2005). In this case, it might be the possibility that the disfluent presentation of the stimuli might make participants more susceptible to post-event information since they were more likely to willing to learn the following information. The second explanation might be disfluent presentation decreased the effectiveness of encoding; therefore, people might not encode the initial information enough, which produces weak traces; therefore, the second information is more likely to interfere with the first information.

The results showed that participants are less accurate for disfluent than fluent stories. However, the integration of the false information from the misinformation phase is not significantly different for fluent and disfluent stories: In other words, the higher probability of choosing information from the misinformation phase compared to the foils was similar for fluent and disfluent conditions. Interestingly, the number of false memories were high for both conditions. It was important to examine the reasons behind the high rate of false memories. Thus, Experiment 2 and Experiment 3 were conducted. In Experiment 2, I controlled the JOLs reactivity effect by excluding the JOL phase from the experiment. In Experiment 3, I controlled for the possible effect of guess responses by inserting a guess-box to the recognition phase and for the possible effect of high number stimuli by subtracting some of the images from the stories. In both Experiment 2 and 3, findings for the number of false memories separately for fluent and disfluent conditions and non-significant effect of fluency on the number of false memories were similar to Experiment 1, suggesting that there was no difference between the fluent and disfluent presentation during encoding in terms of how much misinformation is integrated to the real memories during the misinformation paradigm, regardless of whether the information is presented fluently or disfluently in the original encoding phase.

As it was supported in Experiment 3, there could be some other possible reasons rather than guess responses and noise due to a high number of unrelated details. Firstly, the disfluent presentation might prevent people from encoding the details well enough to recall in the final test. As we can see from the one-sample t-test analysis for the comparison of corrected hit-rates to 0, participants responded below the chance level for disfluent items, indicating that they did not learn well. Secondly, participants might pay more attention to the fluent items to increase their overall

memory performance since they assumed that they would be more likely to remember them (see in Experiment 1). In other words, people might realize disfluent details were not remembered well, so they tried to increase their performance by investing more in fluent ones.

There might also explanations for the result that disfluent presentation did not change the susceptibility to post-event misinformation. It was interesting how people were not more susceptible to misinformation when they did not learn well in the disfluent condition in the initial encoding phase. The first explanation might be related to the effectiveness of metacognitive control processes. Maybe, participants might not try to increase their performance by paying more attention to the following information in the misinformation phase. Even if the misinformation paradigm created a significant amount of false memories, people might not use a further strategy for the disfluent items, so I did not observe more susceptibility. Moreover, participants were informed that they would restudy the same set of stimuli; however, they might still realize the changes in the details; thence, they did not want to learn the wrong information. In a further study, it can be controlled whether participants realized the misinformation manipulation by a question at the end of the experiment so that we can examine whether the awareness changes the results or not. The second explanation might be a high number of false memories for both fluent and disfluent conditions. As the results revealed, participants were susceptible to misinformation in both conditions; therefore, it can be interpreted that disfluent presentation might not have a further effect since the susceptibility to misinformation reached the ceiling.

For further understanding, it was important to compare the findings of Sanchez and Naylor (2018) and our study for false memory. In contrast to Sanchez and Naylor

(2018) findings, we did not observe any difference between fluent and disfluent presentations in terms of false memories. It is important to point out that they found that disfluent presentation increases the number of false recalls. They manipulated perceptual fluency through font types and measured false memories via the DRM paradigm. The nature of the tasks might be the reason behind the difference between the findings. On the one hand, in the DRM paradigm, as we discussed above, people made false memories due to the associations between the given items and the ungiven items in which the semantic relatedness of given items to ungiven items produce false recall of the ungiven item. On the other hand, misinformation manipulation causes false memories because of source-confusion and/or the confusion or omission of initial traces. The disfluent presentation might cause an increase in the confusion due to association; however, it might not be that much affected in the source or trace confusions.

Other critical differences between the nature of the experiments were the way of disfluency manipulation and the types of stimuli. They manipulated the fluency through font size, and the stimuli were the words; however, in our experiments, images were presented via the checker-board mask. Even if the checker-board manipulation was controlled to make it disfluent enough but still perceptually visible, it was still harder than the font size manipulation to see. People might pay attention to the disfluent items due to their low JOLs for those items, which might make them learn the details more effortfully in the disfluent images. Therefore, the possible difference between the fluent and disfluent presentations in terms of the number of memory errors might not be observed.

5.2 Limitations and Further Studies

The higher rate of false memories and the results that are inconsistent with the expectations might have possible explanations. The first reason might be the real absence of the difference, meaning that people might not be open to take the following information more, even if they assume that they would not remember the information in the final test. The second reason might be the effect of instruction on people's strategies. Depending on the findings in the literature, one of the assumptions was that people would prefer restudying and paying more attention to the following information for the items that were presented disfluently in the first encoding phase. However, participants were instructed to pay attention to the information in the misinformation phase because they were deceived that the study's true nature was to examine the effect of repeated study trials on memory performance. Since they were informed about the importance of studying the information in the second encoding (misinformation phase), participants should pay attention to all items similarly in the misinformation phase regardless of the fluency of the first presentation. Therefore, even if they showed a tendency to restudy the items disfluently presented in initial encoding, they might follow the instruction and pay attention to all the details in the second encoding. Thus, misinformation manipulation might produce a similar amount of false memories for fluent and disfluent items. Further studies with different instructions might be conducted to observe whether the disfluent presentation influenced people's learning strategies through JOLs.

Since the false memories are most critical to real criminal cases, the current study aimed to examine the effect of fluency on actual and predicted memory performance and false memory with a method including witnessing an event process; thus, the

stimuli consist of stories with a logical flow. The first limitation was related to the logical-order of stimuli. In most of the metamemory research, participants were mostly asked to provide JOLs for unrelated stimuli (Rhodes & Castel, 2008, 2009); however, in the current project, they provide JOLs for the set of stimuli which creates a story. Stimuli were presented in the same order within stories across participants. Thus, the possible effects of presenting items in a logical order could not be investigated. For instance, participants might use cues from the logical order, meaning that they could guess the response depending on the flow of stories.

Moreover, maybe, the logical order might not just affect the actual memory performance but also influence memory predictions. Since the results showed that participants' JOLs show a continuing pattern from the first presentation to the second presentation, it is essential to examine the further effect of story-like stimuli on JOLs. Participants might have overall judgments for each story depending on their fluency in initial encoding. So, they might remember their initial judgments in the original encoding phase, which influences their process of making JOLs in the misinformation phase. The results might differ in a study where participants make item-by-item JOLs for unrelated items, which might be a further study to observe whether the continuing effect of fluency on JOLs will continue. For instance, images that are unrelated to each other might be fluently and disfluently presented in the initial encoding phase, and then some of their details might be changed for misinformation manipulation. As in Experiment 1 of the current project, they will report JOLs in both encoding phases. In this case, they will make JOLs for each item independently. If the continuing effect in the current project was observed due to the story-like stimuli set, I expected not to see the continuing effect of JOLs from the first presentation to the second presentation.

In daily life, when individuals witness a crime, they might be exposed to misinformation through various sources, such as police-questioning, media, social media, or post-event suggestions. In the court, an eye witness' responsibility is to recall the details of the original event, without being affected by the post-event information. Therefore, in the current study, participants were asked to recall information from the initial encoding phase. This manipulation aimed to use a method closer to the real eye witnessing process in which the witnessing an event followed by misinformation and questioning about the initial event. Thus, I made participants to recall from the specific source (e.g., original encoding phase). It might be interesting to examine whether participants' susceptibility to information changes when they are asked to recall both from initial and second encoding phases.

Moreover, since the images were only used as stimuli, it might be interesting to examine what if we manipulated misinformation through verbal stimuli (e.g., usage of narratives with misinformation, leading questions, etc.), would this change the results? Usage of images might provide participants with the same context in the initial phase, which might increase their overall memory mistakes, especially in fluent condition. If the misinformation is manipulated through verbal stimuli, the rate of susceptibility to false memory might be lower. Another interesting thing might be the investigation of the usage of images in the test phase. The recognition test was presented through images with foil (i.e., unrelated choice), false (i.e., the choice presented in misinformation phase), and correct (i.e., the choice presented in original phase) responses. Using images in the test phase might probably increase overall memory performance. Since the images in the test phase will be the same with either original or misinformation phases, they provide participants with cues that might lead participants to reach an accurate source of information.

The manipulation might be done within or between-subjects design. In a within-subjects design, participants recall half of the stimuli from the initial encoding phase and a half from the second encoding phase, and in between subjects case, half of the participants are asked to recall from the initial encoding and the other half from the second encoding phase. Thus, we can control the effect of retroactive interference, referring that new information interferes with old information (Kreutzer, DeLuca, & Caplan, 2011) so that we can observe the source-confusion due to disfluent presentation. Furthermore, in this kind of method, different combinations (e.g., fluent followed by disfluent, disfluent followed by disfluent) might be used to examine the effect of fluency manipulation on the probability of source confusion.

5.3 Conclusion

To conclude, this study shows that exposure to the post-event information produces false memory for the initial event, and disfluent presentation of the initial event decreases the probability of accurate recall. Moreover, people make accurate predictions for their future memory performances depending on the fluency manipulation: they assume they would be less likely to remember the disfluent information compared to fluent ones, and they remember fluent information more than disfluent information.

These results are important for the real eyewitness testimony cases. If a person witnesses a murder, and s/he partially sees the murderer used a knife, a police officer should be careful about the accuracy of her/his recollection about the event. If the person is exposed to post-event misinformation, then the results suggest that her recollection will probably include false information. As discussed above, one of the

reasons for false memory is the suggestive question with false information (e.g., asking the person was the weapon a gun). One of the factors that experts should consider for eyewitnesses is the perceptual clarity of the eyewitness during the original event. If the crime happened in a poorly-lit area, if there were impediments to clear perception of the events such as physical obstructions, non-corrected eyewitness vision, it is important to evaluate their testimony accordingly. Events that are not clearly perceived will not be remembered well. The current study shows that this type of impediment in the current paradigm does not necessarily increase the integration of false details to the original event, but it increases accurate memories significantly. Finally, the current study implies that eyewitnesses can predict the direction of their subsequent memory performance accurately. If they are unsure of their perceptual clarity during the encoding of the event, they are able to assess this well. Thus, it is a good idea to ask participants whether they saw the event clearly in the first place before using their testimony in the court proceedings.

REFERENCES

Ackil, J. K., & Zaragoza, M. S. (1998). Memorial consequences of forced confabulation: Age differences in susceptibility to false memories. *Developmental Psychology, 34*(6), 1358–1372. doi:10.1037/0012-1649.34.6.1358

Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General, 136*(4), 569–576. doi:10.1037/0096-3445.136.4.569

Arbuckle, T. Y., & Cuddy, L. L. (1969). Discrimination of item strength at time of presentation. *Journal of Experimental Psychology, 81*(1), 126–131. doi:10.1037/h0027455

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge University Press.

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. doi:10.1016/0749-596X(89)90016-8

Begg, I. M., Martin, L. A., & Needham, D. R. (1992). Memory monitoring: How useful is self-knowledge about memory? *European Journal of Cognitive Psychology, 4*(3), 195–218. doi:10.1080/09541449208406182

Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General, 127*(1), 55–68. doi:10.1037/0096-3445.127.1.55

Benjamin, A. S., & Craik, F. I. (2001). Parallel effects of aging and time pressure on memory for source: Evidence from the spacing effect. *Memory & Cognition, 29*(5), 691–697.

Bergman, E. T., & Roediger, H. L. III. (1999). Can Bartlett's repeated reproduction experiments be replicated? *Memory & Cognition, 27*(6), 937–947. doi:10.3758/BF03201224

Besken, M. (2016). Picture-perfect is not perfect for metamemory: Testing the perceptual fluency hypothesis with degraded images. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(9), 1417–1433. doi:10.1037/xlm0000246

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. doi:10.3758/s13421-013-0307-8

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. doi:10.1037/a0034407

Besken, M., Solmaz, E. C., Karaca, M., & Atılgan, N. (2019). Not all perceptual difficulties lower memory predictions: Testing the perceptual fluency hypothesis with rotated and inverted object images. *Memory & Cognition*, *47*(5), 906-922. doi:10.3758/s13421-019-00907-7

Bjork, R. A. (1994). *Memory and metamemory considerations in the training of human beings.* In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (p. 185–205). The MIT Press.

Bornstein, R. F., & D'Agostino, P. R. (1994). The attribution and discounting of perceptual fluency: Preliminary tests of a perceptual fluency/attributional model of the mere exposure effect. *Social Cognition, 12*(2), 103–128. doi: 10.1521/soco.1994.12.2.103

Brainerd, C. J., & Reyna, V. F. (2005). *Oxford psychology series. The science of false memory.* Oxford University Press. doi:10.1093/acprof:oso/9780195154054.001.0001

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*(1), 17–22. doi: 10.1037/h0046671

Double, K. S., Birney, D. P., & Walker, S. A. (2018). A meta-analysis and systematic review of reactivity to judgments of learning. *Memory, 26*(6), 741–750.

Dunlosky, J., & Bjork, R. A. (2008). The integrated nature of metamemory and memory. *Handbook of Metamemory and Memory*, 11–28. doi:10.1080/09658211.2017.1404111

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. doi:10.3758/BF03210921

Dunlosky, J., & Nelson, T. O. (1994). Does the sensitivity of judgments of learning (JOLs) to the effects of various study activities depend on when the JOLs occur. *Journal of Memory and Language, 33*(4), 545-565.

Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self-evaluations undermine students' learning and retention. *Learning and Instruction, 22*(4), 271–280. doi:10.1016/j.learninstruc.2011.08.003

Flavell, J. H., & Wellman, H. M. (1975). Metamemory.

Foster, N. L., & Sahakyan, L. (2012). Metacognition influences item-method directed forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(5), 1309–1324. doi:10.1037/a0027868

Frost, P. (2000). The quality of false memory over time: Is memory for misinformation "remembered" or "known"? *Psychonomic Bulletin & Review, 7*(3), 531–536. doi:10.3758/BF03214367

Gallo, D. (2013). *Associative illusions of memory: False memory research in drm and related tasks*. Psychology Press.

Gallo, D. A., Roberts, M. J., & Seamon, J. G. (1997). Remembering words not presented in lists: Can we avoid creating false memories? *Psychonomic Bulletin & Review, 4*(2), 271–276. doi:10.3758/BF03209405

Goldsmith, M., & Koriat, A. (2007). The strategic regulation of memory accuracy and informativeness. *Psychology of Learning and Motivation*, *48*, 1–60.

Hanczakowski, M., Zawadzka, K., & Cockcroft-McKay, C. (2014). Feeling of knowing and restudy choices. *Psychonomic Bulletin & Review, 21*(6), 1617–1622. doi:10.3758/s13423-014-0619-0

Hege, A. C., & Dodson, C. S. (2004). Why distinctive information reduces false memories: Evidence for both impoverished relational-encoding and distinctiveness heuristic accounts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(4), 787-795. doi:10.1037/0278-7393.30.4.787

Hemmer, P., & Steyvers, M. (2009). A Bayesian account of reconstructive memory. *Topics in Cognitive Science*, *1*(1), 189–202.

Hertzog, C., Dixon, R. A., & Hultsch, D. F. (1990). Relationships between metamemory, memory predictions, and memory task performance in adults. *Psychology and Aging, 5*(2), 215–227. doi: 10.1037/0882-7974.5.2.215

Hertzog, C., Dunlosky, J., Robinson, A. E., & Kidder, D. P. (2003). Encoding fluency is a cue used for judgments about learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(1), 22–34. doi:10.1037/0278-7393.29.1.22

Howe, M. L. (1998). When distinctiveness fails, false memories prevail. *Journal of Experimental Child Psychology, 71*(2), 170–177. doi:10.1006/jecp.1998.2469

Jacoby, L. L., & Kelley, C. M. (1987). Unconscious influences of memory for a prior event. *Personality and Social Psychology Bulletin*, *13*(3), 314–336. doi:10.1177/0146167287133003

Jacoby, L. L., Woloshyn, V., & Kelley, C. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology: General, 118*(2), 115–125. doi:10.1037/0096-3445.118.2.115

James, C. T., Thompson, J. G., & Baldwin, J. M. (1973). The reconstructive process in sentence memory. *Journal of Verbal Learning & Verbal Behavior, 12*(1), 51–63. doi:10.1016/S0022-5371(73)80060-X

Janes, J. L., Rivers, M. L., & Dunlosky, J. (2018). The influence of making judgments of learning on memory performance: Positive, negative, or both? *Psychonomic Bulletin & Review, 25*(6), 2356–2364. doi:10.3758/s13423-018-1463-4.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*(1), 3–28. doi:10.1037/0033-2909.114.

Karpicke, J. D., Butler, A. C., & Roediger III, H. L. (2009). Metacognitive strategies in student learning: Do students practice retrieval when they study on their own? *Memory*, *17*(4), 471–479. doi:10.1080/09658210802647009

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. doi:10.1037/0096-3445.126.4.349

Koriat, A., & Bjork, R. A. (2005). Illusions of competence in monitoring one's knowledge during study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(2), 187–194. doi:10.1037/0278-7393.31.2.187

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

Koriat A., Shitzer-Reichert R. (2002) Metacognitive Judgments and their Accuracy. In: Chambres P., Izaute M., Marescaux PJ. (eds) Metacognition. Springer, Boston, MA. doi: 10.1007/978-1-4615-1099-4_1

Kreutzer, J. S., DeLuca, J., & Caplan, B. (2011). *Encyclopedia of clinical neuropsychology*. SpringerLink (Online service).

Kuhn, D. (2000). Metacognitive development. *Current Directions in Psychological Science*, *9*(5), 178–181. doi:10.1111/1467-8721.00088

Lindsay, D. S., & Johnson, M. K. (1989). The eyewitness suggestibility effect and memory for source. *Memory & Cognition*, *17*(3), 349–358. doi:10.3758/BF03198473

Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology*, *7*(4), 560–572. doi:10.1016/0010-0285(75)90023-7

Loftus, E. F. (1977). Shifting human color memory. *Memory & Cognition*, *5*(6), 696–699. doi:10.3758/BF03197418

Loftus, E. F. (1979). The malleability of human memory: Information introduced after we view an incident can transform memory. *American Scientist*, *67*(3), 312-320.

Loftus, E. F. (2004). Memories of Things Unseen. *Current Directions in Psychological Science*, *13*(4), 145–147. doi:10.1111/j.0963-7214.2004.00294.x

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory*, *4*(1), 19–31. doi:10.1037/0278-7393.4.1.19

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning & Verbal Behavior*, *13*(5), 585–589. doi:10.1016/S0022-5371(74)80011-3

Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric Annals*, *25*(12), 720–725. doi:10.3928/0048-5713-19951201-07

Marsh, R. L., Hicks, J. L., & Cook, G. I. (2005). On the Relationship Between Effort Toward an Ongoing Task and Cue Detection in Event-Based Prospective Memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(1), 68–75. doi:10.1037/0278-7393.31.1.68

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. doi:10.3758/BF03194916

Mazzoni, G., & Nelson, T. O. (1995). Judgments of learning are affected by the kind of encoding in ways that cannot be attributed to the level of recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(5), 1263–1274. doi: 10.1037/0278-7393.21.5.1263

Mccabe, D. P., Presmanes, A. G., Robertson, C. L., & Smith, A. D. (2004). Item-specific processing reduces false memories. *Psychonomic Bulletin & Review*, *11*(6), 1074–1079. doi:10.3758/BF03196739

Mitchell, K. J., & Zaragoza, M. S. (1996). Repeated exposure to suggestion and false memory: The role of contextual variability. *Journal of Memory and Language*, *35*(2), 246–260. doi:10.1006/jmla.1996.0014

Mudd, K., & Govern, J. M. (2004). Conformity to misinformation and time delay negatively affect eyewitness confidence and accuracy. *North American Journal of Psychology*, *6*(2). 227–238.

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. doi:10.1016/j.jml.2013.09.007

Mueller, M.L., Tauber, S.K. & Dunlosky, J. (2013) Contributions of beliefs and processing fluency to the effect of relatedness on judgments of learning. *Psychonomic Bull Rev* 20, 378–384. doi: 10.3758/s13423-012-0343-6

Myers, S. J., Rhodes, M. G., & Hausman, H. E. (2020). Judgments of learning (JOLs) selectively improve memory depending on the type of test. *Memory & Cognition*, 1–14. doi:10.3758/s13421-020-01025-5

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *The Psychology of Learning and Motivation*, *26*, 125–141. doi:10.1016/S0079-7421(08)60053-5

Nomi, J. S., Rhodes, M. G., & Cleary, A. M. (2013). Emotional facial expressions differentially influence predictions and performance for face recognition. *Cognition & emotion*, *27*(1), 141-149. doi:10.1080/02699931.2012.679917

Okado, Y., & Stark, C. (2003). Neural processing associated with true and false memory retrieval. *Cognitive, Affective, & Behavioral Neuroscience*, *3*(4), 323–334. doi:10.3758/CABN.3.4.323

Okado, Y., & Stark, C. E. (2005). Neural activity during encoding predicts false memories created by misinformation. *Learning & Memory*, *12*(1), 3–11. doi:10.1101/lm.87605

Oppenheimer, D. M. (2008). The secret life of fluency. *Trends in Cognitive Sciences*, *12*(6), 237–241. doi:10.1016/j.tics.2008.02.014

Payne, J. D., Nadel, L., Allen, J. J., Thomas, K. G., & Jacobs, W. J. (2002). The effects of experimentally induced stress on false recognition. *Memory*, *10*(1), 1–6. doi: 10.1080/09658210143000119

Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition: An International Journal*, *8*(3), 338–342. doi:10.1006/ccog.1999.0386

Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science*, *9*(1), 45–48. doi:10.1111/1467-9280.00008

Reyna, V. F., & Titcomb, A. L. (1997). Constraints on the suggestibility of eyewitness testimony: A fuzzy-trace theory analysis. *Intersections in Basic and Applied Memory Research*, 157–174.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rhodes, M. G. (2016). Judgments of learning: Methods, data, and theory. *The Oxford handbook of metamemory*, *1*, 65-80. doi:10.1093/oxfordhb/9780199336746.013.4

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. doi:10.1037/a0013684

Rhodes, M. G., & Castel, A. D. (2009). Metacognitive illusions for auditory information: Effects on monitoring and control. *Psychonomic Bulletin & Review*, *16*(3), 550–554. doi:10.3758/PBR.16.3.550

Rhodes, M. G., & Tauber, S. K. (2011). The influence of delaying judgments of learning on metacognitive accuracy: A meta-analytic review. *Psychological Bulletin*, *137*(1), 131. doi:10.1037/a0021705

Roediger III, H. L., & DeSoto, K. A. (2015). Reconstructive memory, psychology of. In N. J. Smelser & B. Baltes (eds.), *International Encyclopedia of the Social and Behavioral Sciences*, 12844–12849. doi: 10.1016/B0-08-043076-7/01521-7

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 803. doi:10.1037/0278-7393.21.4.803

Roediger III, H. L., Jacoby, J. D., & McDermott, K. B. (1996). Misinformation effects in recall: Creating false memories through repeated retrieval. *Journal of Memory and Language*, *35*(2), 300–318. doi:10.1006/jmla.1996.0017

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

Russell, W. A., & Jenkins, J. J. (1954). *Studies on the role of language in behavior: The complete Minnesota norms for responses to 100 words from the Kent-Rosanoff word association test*. Minnesota University Minneapolis.

Sanchez, C. A., & Naylor, J. S. (2018). Disfluent presentations lead to the creation of more false memories. *PLoS One*, *13*(1), Article e0191735. doi:10.1371/journal.pone.0191735

Schaaf, J. M., Alexander, K. W., & Goodman, G. S. (2008). Children's false memory and true disclosure in the face of repeated questions. *Journal of Experimental Child Psychology*, *100*(3), 157–185. doi:10.1016/j.jecp.2007.09.002

Schacter, D. L., Koutstaal, W., & Norman, K. A. (1997). False memories and aging. *Trends in Cognitive Sciences*, *1*(6), 229–236. doi:10.1016/S1364-6613(97)01068-1

Schooler, J. W., Gerhard, D., & Loftus, E. F. (1986). Qualities of the unreal. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(2), 171–181. doi:10.1037/0278-7393.12.2.171

Shah, A. K., & Oppenheimer, D. M. (2008). Heuristics made easy: An effort-reduction framework. *Psychological Bulletin*, *134*(2), 207–222. doi:10.1037/0033-2909.134.2.207

Smelser, N. J., Baltes, P. B., & others. (2001). *International encyclopedia of the social & behavioral sciences* (Vol. 11). Elsevier Amsterdam.

Soderstrom, N. C., Clark, C. T., Halamish, V., & Bjork, E. L. (2015). Judgments of learning as memory modifiers. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(2), 553–558. doi:10.1037/a0038388

Son, L. K., & Metcalfe, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(1), 204–221. doi:10.1037/0278-7393.26.1.204

Song, H., & Schwarz, N. (2008). If it's hard to read, it's hard to do: Processing fluency affects effort prediction and motivation. *Psychological Science, 19*(10), 986–988. doi:10.1111/j.1467-9280.2008.02189.x

Stark, C. E., Okado, Y., & Loftus, E. F. (2010). Imaging the reconstruction of true and false memories using sensory reactivation and the misinformation paradigms. *Learning & Memory*, *17*(10), 485–488. doi:10.1101/lm.1845710

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. doi:10.3758/s13423-011-0114-9

Susser, J. A., Mulligan, N. W., & Besken, M. (2013). The effects of list composition and perceptual fluency on judgments of learning (JOLs). *Memory & Cognition*, *41*(7), 1000–1011. doi:10.3758/s13421-013-0323-8

Tauber, S. K., & Rhodes, M. G. (2012). Measuring memory monitoring with judgements of retention (JORs). *The Quarterly Journal of Experimental Psychology*, *65*(7), 1376–1396. doi:10.1080/17470218.2012.656665

The Jamovi Project (2019). *jamovi*. (Version 1.1) [Computer Software]. Retrieved from https://www.jamovi.org.

Tiede, H. L., Derksen, C., & Leboe, J. P. (2009). An investigation of increases in metamemory confidence across multiple study trials. *Memory*, *17*(3), 288–300. doi:10.1080/09658210802665837

Tyng, C. M., Amin, H. U., Saad, M. N. M., & Malik, A. S. (2017). The influences of emotion on learning and memory. *Frontiers in Psychology*, *8*, Article 1454. doi:10.3389/fpsyg.2017.01454

Undorf, M., & Ackerman, R. (2017). The puzzle of study time allocation for the most challenging items. *Psychonomic Bulletin & Review*, *24*(6), 2003–2011. doi:10.3758/s13423-017-1261-4

Undorf, M., & Zimdahl, M. F. (2019). Metamemory and memory for a wide range of font sizes: What is the contribution of perceptual fluency? *Journal of*

Experimental Psychology: Learning, Memory, and Cognition, 45(1), 97–109. doi:10.1037/xlm0000571

Weinstein, Y., & Shanks, D. R. (2010). Rapid induction of false memory for pictures. *Memory, 18*(5), 533–542. doi:10.1080/09658211.2010.483232

Witherby, A. E., & Tauber, S. K. (2017). The influence of judgments of learning on long-term learning and short-term performance. *Journal of Applied Research in Memory and Cognition, 6*(4), 496–503. doi:10.1016/j.jarmac.2017.08.004

Wolk, D. A., Schacter, D. L., Berman, A. R., Holcomb, P. J., Daffner, K. R., & Budson, A. E. (2004). An electrophysiological investigation of the relationship between conceptual fluency and familiarity. *Neuroscience Letters, 369*(2), 150–155.doi:10.1016/j.neulet.2004.07.081

Wong, S. S. H., & Lim, S. W. H. (2019). From JOLs to JOLs+: Directing learners' attention in retrieval practice to boost integrative argumentation. *Journal of Experimental Psychology: Applied, 25*(4), 543–557. doi:10.1037/xap0000225

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. doi:10.3758/s13421-012-0255-8

Zhu, B., Chen, C., Loftus, E. F., Lin, C., He, Q., Chen, C., … Dong, Q. (2010). Individual differences in false memory from misinformation: Cognitive factors. *Memory, 18*(5), 543–555. doi:10.1080/09658211.2010.487051

Zimmerman, C. A., & Kelley, C. M. (2010). "I'll remember this!" Effects of emotionality on memory predictions versus memory performance. *Journal of Memory and Language, 62*(3), 240-253. doi:10.1016/j.jml.2009.11.004

APPENDICES

APPENDIX A EXAMPLES OF CHANGES IN THE IMAGES

APPENDIX B FORMULAS FOR ACCURATE AND FALSE MEMORIES

Abbreviations

False Memory: FaMe

False Alarm: FA

Accuracy: Ac

T: Total

F: Fluency

DF: Disfluency

Foil: false and unrelated choices in the recognition test (e.g., crowbar)

Misinformation: 'mis' – false and the choices presented differently in the second presentation (e.g., hanger)

Correct: correct choices (e.g., credit card)

Control = 'cont' – the stimuli which do not change in the second presentation

Critical = 'crit' – the stimuli which change in the second presentation

Proportion = Prop

$$TAc = \frac{The\ number\ of\ correct\ responses}{The\ number\ of\ questions\ (144)}$$

$$TAc_F = \frac{The\ number\ of\ correct\ responses\ in\ fluent\ condition}{The\ number\ of\ questions\ in\ fluent\ condition\ (72)}$$

$$Ac_F_Crit = \frac{Number\ of\ correct\ responses\ to\ questions\ with\ mis\ in\ F\ condition}{Number\ of\ questions\ with\ mis\ in\ F\ condition\ (48)}$$

$$Ac_F_Cont = \frac{Number\ of\ correct\ responses\ to\ questions\ with\ foils\ in\ F\ condition}{Number\ of\ questions\ with\ foils\ in\ F\ condition\ (24)}$$

$$TAc_DF = \frac{The\ number\ of\ correct\ responses\ in\ disfluent\ condition}{The\ number\ of\ questions\ in\ disfluent\ condition\ (72)}$$

$$Ac_DF_Crit = \frac{Number\ of\ correct\ responses\ to\ questions\ with\ mis\ in\ DF\ condition}{Number\ of\ questions\ with\ mis\ in\ DF\ condition}\ (48)$$

$$Ac_DF_Cont = \frac{Number\ of\ correct\ responses\ to\ questions\ with\ foils\ in\ DF\ condition}{Number\ of\ questions\ with\ foils\ in\ DF\ condition}\ (24)$$

$$T_FaMe_Prop = \frac{Number\ of\ mis\ responses}{Sum\ of\ mis\ and\ foil\ responses}$$

$$FaMe_Prop_F = \frac{Number\ of\ mis\ responses\ in\ fluent\ condition}{Sum\ of\ mis\ and\ foil\ responses\ in\ fluent\ condition}$$

$$FaMe_Prop_DF = \frac{Number\ of\ mis\ responses\ in\ disfluent\ condition}{Sum\ of\ mis\ and\ foil\ responses\ in\ disfluent\ condition}$$

$$Foil_Prop_F = \frac{Number\ of\ foil\ responses\ in\ fluent\ condition}{Sum\ of\ mis\ and\ foil\ responses\ in\ fluent\ condition}$$

$$Foil_Prop_DF = \frac{Number\ of\ foil\ responses\ in\ disfluent\ condition}{Sum\ of\ mis\ and\ foil\ responses\ in\ disfluent\ condition}$$

APPENDIX C ORDER OF CONDITIONS – PARTIAL APPLICATION OF LATIN SQUARE COUNTERBALANCE

a. Order of Conditions in Experiment 1 & Experiment 2 (S: Story, DF: Disfluent, F: Fluent, Numbers refers to the stories (e.g., S1 means first story)

Conditions	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6	Trial 7	Trial 8
C1	S1-DF	S2-F	S8-DF	S3-F	S7-F	S4-DF	S6-DF	S5-F
C2	S2-DF	S3-F	S1-F	S5-DF	S8-DF	S4-F	S7-F	S6-DF
C3	S3-DF	S4-DF	S2-F	S5-F	S1-DF	S6-DF	S8-F	S7-F
C4	S4-DF	S5-DF	S3-F	S7-DF	S1-F	S6-F	S2-DF	S8-F
C5	S5-F	S7-DF	S4-F	S2-DF	S8-DF	S3-F	S6-F	S1-DF
C6	S6-F	S7-DF	S5-DF	S8-F	S4-F	S1-DF	S3-DF	S2-F
C7	S7-F	S6-F	S8-DF	S2-DF	S5-F	S1-F	S4-DF	S3-DF
C8	S8-F	S1-F	S7-DF	S2-F	S6-DF	S3-DF	S5-DF	S4-F

b. Order of Conditions in Experiment 3 (S: Story, DF: Disfluent, F: Fluent)

Conditions	Trial 1	Trial 2	Trial 3	Trial 4
C1	S1-DF	S2-DF	S4-F	S3-F
C2	S2-F	S3-DF	S1-DF	S4-F
C3	S3-F	S4-DF	S2-F	S1-DF
C4	S4-DF	S1-F	S3-F	S2-DF
C5	S4-DF	S2-F	S3-DF	S1-F
C6	S1-F	S4-F	S2-DF	S3-DF

APPENDIX D COMPUTER SCREEN FOR FIRST INSTRUCTION

Deneyimize hoşgeldiniz.

Bu deney bellek performansını etkileyen faktörleri belirlemek için hazırlanmıştır. Deneyimiz 4 bölümden oluşmaktadır; çalışma, tekrar çalışma, sudoku ve test

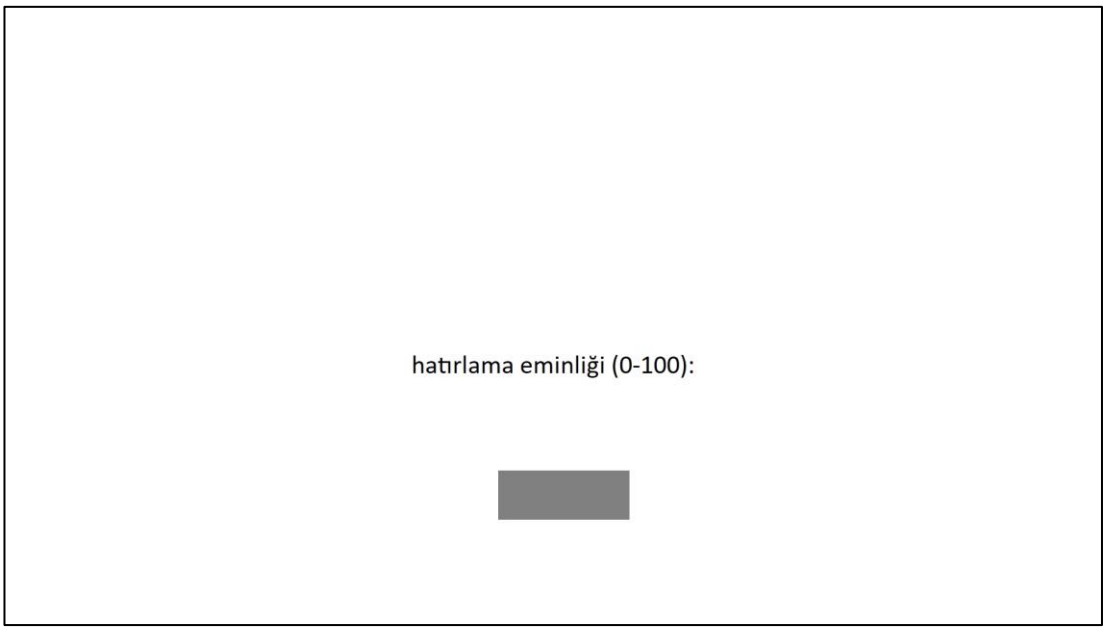
Çalışma Bölümü

Çalışma kısmında size 8 adet hikaye gösterilecektir. Bu hikayeler size resim olarak sunulacaktır.Bu hikayelerin bir kısmı net olmayan (dama tahtasıyla üzeri kapatılmış şekilde) bir kısmı ise net (herhangi bir engel olmadan) sunulacaktır. Sizden istenen her bir resmi dikkatli bir şekilde incelemeniz ve detaylara dikkat etmenizdir. Her bir resimden sonra görmüş olduğunuz resmi ileride ne kadar iyi hatırlayacağınızı değerlendirmeniz istenecektir. Hatırlama eminliğinizi 1'den 100'e kadar olan sayıları kullanarak bildirmeniz beklenmektedir. Değerlendirmenizi 1'den 100'e kadar olan sayıların tamamını düşünerek ve mümkün olduğunca tüm aralığı kullanarak belirtiniz. Hatırlama eminliğinizi belirtip enter tuşuna basınız.

Herhangi bir sorunuz varsa lütfen başlamadan sorunuz. Deney süresince herhangi bir soru sormamanız beklenmektedir.

Hazırsanız başlamak için boşluk (space) tuşuna basınız.

APPENDIX E COMPUTER SCREEN VIEW OF AN IMAGE AND JOLs



APPENDIX F COMPUTER SCREEN FOR RESTUDY PHASE INSTRUCTION

Çalışma kısmının sonuna geldiniz.

Şimdi yeniden çalışma bölümüne geçeceksiniz. Bu bölümde size birinci kısımda sunulan hikayeler yeniden ve hepsi açık/net (herhangi bir engel olmadan)sunulacaktır. Bu kısımda birinci kısımdaki gibi resimlerin detaylarını dikkatli bir şekilde incelemeniz beklenmektedir. Bu kısımda da her resimden sonra hatırlama eminliğini birinci kısımla aynı olacak şekilde belirtmeniz beklenmektedir.

Hazırsanız başlamak için boşluk (space) tuşuna basınız.

APPENDIX G COMPUTER SCREEN FOR TEST PHASE INSTRUCTION

Test Bölümü

Bu bölümde size hikayelerin detaylarıyla ilgili sorular yöneltilecektir. Her hikaye için 18 test sorusu cevaplamanız beklenmektedir. Size her hikayenin testinden önce, testin hangi hikayeyle ilgili oldugu hatırlatılacaktır. Sizden beklenen cevaplarınızı hikayelerin birinci kısımda sunuldukları şeklini düşünerek vermeniz.