

T.C.
BAHCESEHIR UNIVERSITY
GRADUATE SCHOOL
COGNITIVE NEUROPSYCHOLOGY HEAD OF THE DEPARTMENT

**INVESTIGATION METACOGNITIVE JUDGMENTS UNDER
MISINFORMATION EFFECT IN THE RETRIEVAL-INDUCED
FORGETTING PARADIGM: EFFECTS OF CUE AND FEEDBACK**

MASTER'S THESIS

CEREN AKCAM

ISTANBUL 2024

T.C.
BAHCESEHIR UNIVERSITY
GRADUATE SCHOOL
COGNITIVE NEUROPSYCHOLOGY HEAD OF THE DEPARTMENT

**INVESTIGATION METACOGNITIVE JUDGMENTS UNDER
MISINFORMATION EFFECT IN THE RETRIEVAL-INDUCED
FORGETTING PARADIGM: EFFECTS OF CUE AND FEEDBACK**

MASTER'S THESIS

THESIS ADVISOR
PROF. DR. METEHAN IRAK

ISTANBUL 2024



T.C.
BAHCESEHIR UNIVERSITY
GRADUATE SCHOOL

MASTER THESIS APPROVAL FORM

Program Name:	Cognitive Neuropsychology (English)
Student's Name and Surname:	Ceren Akçam
Name Of The Thesis:	Investigation Metacognitive Judgments Under Misinformation Effect in Retrieval-Induced Forgetting Paradigm: Effects of Cue and Feedback
Thesis Defense Date:	10.07.2024

This thesis has been approved by the Graduate School which has fulfilled the necessary conditions as Master thesis.

Doç. Dr. Yücel Batu SALMAN
Institute Director

This thesis was read by us, quality and content as a Master's thesis has been seen and accepted as sufficient.

	Title/Name	Institution	Signature
Thesis Advisor's	Prof. Dr. Metehan Irak	Bahçeşehir University	
Member's	Dr. Öğr. Gör. Hale Ögel Balaban	Bahçeşehir University	
Member's	Prof. Dr. Murat Kurt	Acıbadem University	



I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: Ceren Akçam

Signature:

ABSTRACT

INVESTIGATION METACOGNITIVE JUDGMENTS UNDER MISINFORMATION EFFECT IN THE RETRIEVAL-INDUCED FORGETTING PARADIGM: EFFECTS OF CUE AND FEEDBACK

Ceren Akçam

Master's Program in Cognitive Neuropsychology

Supervisor: Prof. Dr. Metehan Irak

June 2024 , 51 pages

Although memory distortions and forgetting are highly studied cognitive phenomena, the literature lacks a comprehensive explanation of the act of metacognitive judgments related to those misconceptions of memory. Thus, the current study explores the interplay between metacognitive judgments and memory fallibility during a retrieval-induced forgetting (RIF) task including the misinformation effect in 132 participants. For this task, participants engaged in the retrieval practice of stolen items from a simulated burglary scenario with subsequent recall tasks (with or without cues) including the acquisition of JOL judgments followed by exposure to misinformation and recognition tasks requiring prior acquisition of FOK judgments. Results confirm the presence of the RIF effect and its persistence during the misinformation effect. Moreover, cues were found to be effective in decreasing JOL judgments while not effective in FOK judgments. However, results indicate higher FOK judgments for items associated with retrieval practice. Stressing the influence of retrieval cues and practice on metacognitive assessments. Notably, results showed higher JOL and FOK judgments for correct answers compared to incorrect ones implying successful metacognitive monitoring to some extent. Results also indicate higher relative accuracy of FOK judgments in test conditions where cues are present/ feedback absent. Furthermore, an increase in

overall memory judgments was observed after the retrieval practice while fallacious feedback did not impact the overall memory judgment taken at the end of the study. The results are further discussed in context with the relevant literature.

Keywords: Metacognitive Judgments, Misinformation Effect, Retrieval-Induced Forgetting, Cues, Fallacious Feedback.



ÖZ

YANLIŞ BİLGİ ETKİSİNİ İÇEREN GERİ GETİRME YOLLU UNUTMA PARADİGMASINDA ÜST BİLİŞSEL YARGILARIN İNCELENMESİ: İPUCU VE GERİ BİLDİRİMİN ETKİSİ

Ceren Akçam

Bilişsel Nöropsikoloji Yüksek Lisans Programı

Tez Danışmanı: Prof. Dr. Metehan Irak

Haziran 2024 , 51 sayfa

Bellek bozulmaları ve unutma, yoğun bir şekilde çalışılan bilişsel süreçler olmasına rağmen, bellek bozulmaları ve unutma ile ilgili üst bilişsel değerlendirmeler yeterince araştırılmamıştır. Bu amaçla, bu çalışmada üst bilişsel yargılar ve bellek yanılgıları arasındaki ilişkiler, yanlış bilgi etkisi içeren geri getirme yollu unutma görevi kullanılarak 132 katılımcı ile incelenmiştir. Bu çalışma için katılımcılar, önce simüle edilmiş bir hırsızlık senaryosunu okuyup öğrendiler. Ardından da bununla ilgili öğrenme kararlarının değerlendirilmesini içeren hatırlama (ipucuyla veya ipuçsuz) görevini, yanlış bilgi etkisinin yer aldığı ve bilme hissi yargılarının değerlendirildiği tanıma görevini yerine getirdiler. Çalışmadan elde edilen sonuçlar geri getirme yollu unutma etkisinin varlığını ve yanlış bilgi etkisinden sonraki devamlılığını ortaya koymuştur. Ayrıca, ipuçlarının katılımcıların öğrenme kararlarını düşürmede etkili olduğu bulunurken, bunun bilme hissi değerlendirmelerini etkilemediği gözlenmiştir. Buna karşın sonuçlar, geri getirme pratiği ile ilişkilendirilen öğeler için daha yüksek bilme hissi değerlendirmeleri olduğunu göstermiştir. Bu sonuç, geri getirmede ipucu ve pratiğin üst bilişsel değerlendirmeler üzerindeki etkisine işaret etmiştir. Özellikle, doğru ve yanlış cevaplar için üst bilişsel değerlendirmeler karşılaştırıldığında, sonuçlar doğru cevaplar için daha yüksek bilme hissi ve öğrenme kararları değerlendirmeleri

ortaya koyarak üst bilişsel izlemenin belirli bir düzeye kadar başarılı olduğunu göstermiştir. Ayrıca sonuçlar, ipucunun mevcut olduğu / geri bildirimin olmadığı test koşullarında bilme hissi yargılarının göreceli doğruluğunun daha yüksek olduğunu göstermiştir. Son olarak, geri getirme pratiğinden sonra genel bellek yargılarında artış gözlemlenirken yanlış geri bildirimin çalışmanın sonunda yer alan genel bellek yargısını etkilemediği görülmüştür. Sonuçlar, ilgili literatür bağlamında tartışılmıştır.

Anahtar Kelimeler: Üst Bilişsel Yargılar, Yanlış Bilgi Etkisi, Geri Getirme Yollu Unutma, İpucu, Yanlış Geri Bildirim.



*To My Dear Family,
I Love You to the Moon and Back...*

ACKNOWLEDGEMENTS

I would first like to express my deepest gratitude to my thesis advisor Prof. Dr. Metehan Irak for his unwavering support and guidance. His patience and encouragement were invaluable throughout my studies. He opened countless windows to new perspectives and patiently guided me, helping me grow and learn. Metaphorically, the road was long and arduous, but he removed many stones from my path, making the journey more navigable. His expertise and mentorship have been instrumental in the completion of this study.

I would also like to extend my heartfelt thanks to my jury members, Dr. Hale Ögel Balaban and Prof. Dr. Murat Kurt for their invaluable contributions and insightful feedback. Their support and warm smiles made my defense process a positive and enriching experience.

My deepest, heartfelt gratitude goes to my dear family, Ahmet and Nurcan Akçam. They shared the challenges with me and supported me throughout this journey with their boundless love. I always felt their supportive hands on my shoulders. They always encouraged me to take one step further. Their unwavering presence, standing beside me with every decision made this path smoother and more manageable.

I am also deeply thankful to my beloved boyfriend, Efe Uslu, who always believed in me and supported me, doing whatever he could to help me. Especially, during the data collection and data entry processes, he worked tirelessly to assist me. He shared my anxieties, listened to my complaints, and encouraged me to overcome my emotional burnout. My family and my boyfriend made this journey much bearable, and their relentless support was my constant source of strength.

Lastly, I owe a special thanks to Dr. Murat Gökçe Üngör, who provided his voluntary support during the computerization of the experimental task used in this study. He brought his expertise and unconditional support to the table, for which I am deeply grateful.

TABLE OF CONTENTS

ETHICAL CONDUCT	iii
ABSTRACT	iv
ÖZ	vi
DEDICATION.....	vii
ACKNOWLEDGEMENTS.....	ix
LIST OF TABLES.....	xiii
LIST OF FIGURES	xiv
Chapter 1: Introduction.....	1
1.1 Bases of Memory and Cognition	1
1.2 Metacognition	2
1.3 Metacognitive Judgments	4
1.3.1 Feeling of Knowing	5
1.3.2 Judgment of Learning.....	6
1.4 Memory Fallibility.....	7
1.5 Metacognition and Memory Fallibility.....	9
1.6 Memory Performance with Cues and Feedback.....	11
1.7 The Current Study	12
Chapter 2: Method	15
2.1 Participants.	15
2.2 Materials	15
2.2.1 Form of Sociodemographic Information.....	15
2.2.2 Beck Anxiety Inventory.	16

2.2.3 Beck Depression Inventory.	16
2.2.4 The Retrieval-Induced Forgetting and Misinformation Task.	16
2.2.5 Study Phase.	17
2.2.6 Recall Phase.	17
2.2.7 Misinformation Task.	17
2.2.8 Forced Choice Recognition Task.	17
2.2.9 Metacognitive Judgment Rating.....	17
2.2.10 Rating of Subjective Confidence Judgment.	19
2.3 Procedure	19
Chapter 3: Results.....	22
3.1 Results for Recall Performance	22
3.2 Results for Recognition Performance.....	25
3.3 Results for Metacognitive Judgments.....	28
3.3.1 Results for JOL Judgments.....	28
3.3.2 Results for FOK Judgments	32
3.3.3 Results for Overall Memory Confidence	36
3.4 Correlations Among Metacognitive Judgments	37
Chapter 4 : Discussion	39
4.1 Summary of the Results.....	39
4.2 Discussion of the Results for Recall Performance	40
4.3 Discussion of the Results for Recognition Performance	41
4.4 Discussion of the Results for Metacognitive Judgments	42
4.5 Strengths and Limitations	50
4.6. Contributions to the Field and Recommendations for Future Research	50
REFERENCES	52

APPENDICES	59
A. The Informed Consent Form	60
B. Form of Sociodemographic Information	62
C. Beck Anxiety Inventory (BAI)	64
D. Beck Depression Inventory (BDI)	65
E. Study Materials	66
F. Retrieval Practice Questions	70
G. Misinformation Questions	72



LIST OF TABLES

TABLES

Table 1 Demographics of the Participants on a Group Basis.....	23
Table 2 Difference of Item Categories on Recall. Repeated Measure ANOVA Results With Descriptive Statistics	24
Table 3 Effects of Testing Condition on Recall Performance. One-Way ANOVA Results With Descriptive Statistics	25
Table 4 Difference of Item Categories on Recognition. Repeated Measure ANOVA Results With Descriptive Statistics	27
Table 5 Effects of Testing Condition on Recognition Performance. One-Way ANOVA Results With Descriptive Statistics	28
Table 6 Difference of Item Categories on JOL Judgments. Repeated Measures ANOVA Results With Descriptive Statistics.....	29
Table 7 Effects of Testing Condition on JOL Judgments. One-Way ANOVA Results With Descriptive Statistics.....	30
Table 8 Effects of Testing Condition on JOL Accuracy. One-Way ANOVA Results With Descriptive Statistics.....	32
Table 9 Difference of Item Categories on FOK Judgments. Repeated Measures ANOVA Results With Descriptive Statistics.....	33
Table 10 Effects of Testing Condition on FOK Judgments. One-Way ANOVA Results With Descriptive Statistics	34
Table 11 Effects of Testing Condition on FOK Accuracy. One-Way ANOVA Results with Descriptive Statistics.....	35
Table 12 Mean Comparison of Retrospective Memory Judgments After and Before Retrieval Practice	36
Table 13 Effects of Testing Condition Overall Memory Judgments. One-Way ANOVA Results with Descriptive Statistics.....	37
Table 14 Correlations Among Metacognitive Judgments	38

LIST OF FIGURES

FIGURES

Figure 1. A Schematic Depiction of the Task Was Used in the Study 18



Chapter 1

Introduction

1.1 Bases of Memory and Cognition

Human cognition refers to the brain's capability of information storage, retrieval, recognition, and processing (Khera & Rangasamy, 2021). It also involves mental capabilities containing problem-solving, remembering, reasoning, abstract thinking, information comprehension, and learning (Gottfredson, 1997). Since all these processes are named under the broader term cognition, it is defined as an advanced meta-processing mechanism in which various mental processes function together and it governs all mental activities (Irak, 2019). Memory includes storing and recalling information according to the current need (Ashcraft & Radvansky, 2018). Ashcraft and Radvansky (2018) pointed out that the retrieval process of memory refers to both the intentional (conscious) and automatic retrieval of the information. Thus, the term encapsulates the mental processes to acquire and retain the information for subsequent use (Ashcraft & Radvansky, 2018). Melton (1963) defines mental activities that are referred to by memory as encoding (known as acquisition or learning), retention, and retrieval. On the other hand, it is known that memory comprises several interconnected systems that encompass neural substrates and their associated behavioral and cognitive counterparts (Tulving, 1984).

Sutton et al. (2010) proposed the evolution of memory research, particularly after the case of brain-injured patient HM. This led to a distinction between declarative (conscious recollection of events and facts) and non-declarative (procedural) memory by Squire (1992). Non-declarative memory includes skills learned through actions, habit formation and emotional learning. Another distinction within declarative memory is semantic (general knowledge) and episodic (personal experiences) memory. According to the hierarchical organization of memory model by Tulving (1984), procedural memory at the lowest level contains semantic memory, which in turn contains episodic memory. Each type of memory has different functions, with procedural memory sustaining learned associations, semantic memory constructing mental models, and episodic memory storing personally experienced events and

temporal relations. Tulving also discussed the different types of consciousness associated with each memory system: procedural memory with non-knowing consciousness, semantic memory with knowing consciousness, and episodic memory with self-knowing consciousness. Sutton et al. (2010) also discussed that the self-memory system model discrete people's autobiographical knowledge as; lifetime periods, general events, and event-specific details. Hence, autobiographical memories are formed through stable activation within these knowledge patterns. Additionally, Conway et al. (2019) highlighted that in long-term memory, autobiographical memory knowledge structures systematize episodic memories and information with temporary activation patterns that consistently emerge and disperse, giving rise to the memories caused by cues. Control processes within the SMS model, driven by the working self, generate unconscious mental models to achieve goals (Conway et al., 2004), facilitating or inhibiting target memories based on current goals (Sutton et al., 2010). Thus, memory assessments related to episodic memory typically involve learning, recall, and recognition stages. In recall tasks, participants are expected to remember previously learned information, while recognition tasks require identifying studied materials among distractors. These stages remain prominent in the memory research literature, reflecting how individuals base critical decisions on their memory knowledge. Students, for instance, encode information through repetition (which Ashcraft and Radvansky (2018) stress to be the key to storing episodic information), assess their learning, and recognize it during exams. Memory, like any functioning system, has its pitfalls and distortions, emphasizing the importance of laboratory studies in evaluating memory accuracy, understanding pitfalls, and enhancing memory. Further laboratory studies concerning memory and cognition, lead to research questions regarding the existence of a mechanism governing one's cognitive actions and how this mechanism operates which leads to the proposal of the concept of metacognition.

1.2 Metacognition

The concept of metacognition emerged as an idea of a system that inspects cognitive actions (Flavell, 1979). Hence, metacognition can be defined as the awareness of one's thought processes and the ability to think about them (Dunlosky &

Metcalfe, 2009). Nelson and Narens (1990) explain the relationship between metacognition and cognition in two related frameworks. While the object level represents the ongoing cognitive activity the meta-level explains a modal of individuals understanding of the ongoing cognitive activity. Nelson and Narens (1990) suggested in their framework that people's awareness of their thought processes occurs when the meta-level represents the incoming information from the object-level thought processes. Although metacognitive acts seem difficult to understand, they are actions that we use frequently in daily life. For instance, individuals may have a problem remembering the names of the individuals that they have met or where they met them. Thus, they may record the names of individuals whom they have met and where they met along with their phone numbers. Realizing and remembering this forgetfulness and developing strategies against this is an example of a metacognitive process. Furthermore, the concept of metacognition is a multifaceted one, encompassing metacognitive knowledge, monitoring, and control (Dunlosky & Metcalfe, 2009).

Pintrich (2002) stated that metacognitive knowledge is related to an individual's cognition about declarative knowledge which is composed of episodes, beliefs, and facts that can be recalled from long-term memory. Also, Squire (1986) noted that they are accessible to conscious awareness. Metacognitive knowledge includes both general and more specific facts and beliefs about cognition (Pintrich, 2002). On the other hand, evaluating an ongoing process or current state of cognitive activity is defined as metacognitive monitoring. In the experimental studies, researchers asked participants to explicitly judge a cognitive state to assess metacognitive monitoring. However, the ability to judge does not always indicate these judgments are accurate. There is a possibility of those beliefs and judgments being inaccurate. Therewithal, regulation of ongoing activity (stopping it, changing it, continuing it) depends on metacognitive control (Dunlosky & Metcalfe, 2009). Additionally, Nelson and Narens (1994) propose that information from the meta-level is utilized to influence ongoing activity at the object level. Therefore, he argues that metacognitive control is applied whenever the object level is modified by the meta-level (Nelson & Narens, 1994).

Furthermore, an individual's understanding of how his or her memory works or fails is defined as metamemory (Ashcraft & Radvansky, 2018). Nonetheless, memory is a form of cognition, and since metacognition inspects cognitive actions, it takes place under it. Also, Irak et al. (2024) stress that since metacognition is predominantly

explored by focusing on the memory processes, the concept denoting the cognitive operations applied to memory is termed meta-memory. However, by showing overconfidence or underconfidence about subsequently learned material, individuals may be misled by metamemory (Koriat et al., 2002). To sum up, in literature, the awareness of one's cognition and memory is discussed under the terms of metacognition and metamemory. All these discussions have led to the formation of the discussion of how individuals make assessments of their cognitive abilities, which is discussed under the term of metacognitive judgments.

1.3 Metacognitive Judgments

Metamemory judgments can be retrospective or prospective (Irak et al., 2023). Nelson and Narens (1990) indicate that while retrospective monitoring includes a person's confidence judgment of a previous recall response, prospective monitoring relates to judgments about subsequent responses. Thus, while retrospective monitoring is associated with recognition and recall processes in cognition, prospective monitoring begins after the retrospective monitoring process is ended (Irak, 2005). Prospective monitoring is divided into three subcategories by Nelson and Narens (1990); Ease of Learning (EOL), Judgment of Learning (JOL), and Feeling of Knowing (FOK). According to Nelson and Narens (1990), While, Ease of learning (EOL) judgments include predictions about whether material that has not yet been learned will be difficult/easy to learn, JOL occurs during and after learning and is based on predicting future performance depending on items that can still be remembered. Also, the FOK judgments include judgments about the probability that an item, that is not currently remembered, will be remembered in the future (Nelson and Narens, 1990).

The basic research method in memory involves acquiring (study) and retrieving the presented items. In the assessment of metamemory, researchers generally ask participants to judge their memory and retrieval processes. To measure the metamemory judgments, such studies assess global (e.g. see Connor & Dunlosky, 1997) and item-by-item predictions. While global predictions are an estimation of the number of items that participants will be able to retrieve correctly during the testing phase, which can be made before or after the study phase, item-by-item judgments are

made individually per item. For example, item-by-item JOLs are based on each item recently studied; individuals then make item-by-item predictions about the likelihood of being remembered for that item at a later time (Zhou et al., 2016). Thus, while participants make this judgment immediately after studying each item this is called immediate JOLs, judgments that were made item by item after participants studied all of the items were called delayed JOLs (Zhou et al., 2016). It is also discussed that there are a variety of metacognitive judgments that occur at different stages of the learning or retrieval process, and their potential impact on the control of these processes varies. For instance, participants might report their confidence about the possibility of a response being remembered in the future by making a JOL judgment, and later in an experiment, participants may make an FOK judgment by reporting their likelihood of recognizing the studied item at a later time which in the following section, FOK and JOL judgments will be discussed.

1.3.1 Feeling of knowing. Dunlosky and Metcalfe (2009) indicated that the first experimental footprint about metacognition was made about the FOK when Josep Hart developed the recall-judgment-recognition paradigm. In this paradigm firstly participants answered a set of questions. Later, for the questions that were not answered or incorrectly answered participants were asked to evaluate their likelihood of recognizing the answer in a multiple-choice task. After making the FOK judgments about those questions, they tried to recognize the correct answer between the choices (Hart, 1966). Also, Dunlosky and Metcalfe (2009) indicated that Hart's conclusion was that, even though participants were unable to come up with the right answer, they were able to correctly recognize it. It is presented as a prove that even if the recall does not succeed, the answer is still in the memory.

Two main hypotheses have been proposed to explain FOK judgments. The target-accessibility hypothesis (Koriat, 1993) suggests that people retrieve partial information about a target item that they cannot currently recall and that they use this information when giving FOK ratings, while the cue-familiarity hypothesis (Schwartz & Metcalfe, 1992) states that FOK ratings are based on a sense of familiarity (Irak et al, 2019).

1.3.2 Judgment of learning. A person's assessment related to how well an item is learned and their prediction of the probability that they will remember that item in the future when the cue is provided is called JOL judgments (Son & Metcalfe, 2005). As noted by Rhodes (2016), Arbuckle and Cuddy's pioneering work in 1969 was the first study to assess JOLs and methodology that has minimally changed since. They asked participants to study word-word and word-number pairs and made yes/ no judgments about their future memory performance and later reported their future memory performance on a point Likert scale. The study concluded that items with 'yes' answers and higher ratings were more likely to be recalled.

Subsequent studies indicate that highly emotional content, concrete rather than abstract items, faces, large font words, and louder words tend to produce higher JOLs. Also, these studies brought to mind the importance of evaluating the accuracy of metacognitive judgments. For assessing the accuracy, Nelson (1984) proposed using a gamma correlation ranging from -1.00 to +1.00. High JOLs for items likely to be remembered and low JOLs for items unlikely to be remembered lead to a positive gamma correlation, indicating better relative accuracy. However, the accuracy of JOLs depends on some factors. For instance, Nelson (2001) indicated that the accuracy of JOLs after a short delay was higher compared to JOLs that were given immediately after the study. It is also stressed that intentional learning yields more accurate JOLs (Mazzonni & Nelson, 1995), and timing, practice, and testing also affect the JOL performance and relative accuracy (Rhodes, 2016). Yet, there are some highlighted contradictions in the literature. For instance, although test experience boosts relative accuracy Koriat et al. (2002) found that the underconfidence is reported for subsequent trials compared to the first trials (underconfidence with practice effect). Later, Dunlosky and Metcalfe (2009) indicated that underconfidence with practice effect increases when participants are given feedback about the correctness of their answers.

Two main hypotheses have been proposed to explain JOL judgments. The retrieval fluency hypothesis suggests rapid access to more information regarding a desired response can increase its confidence in recollection of it. Also, the cue utilization approach suggests that various cues are used to make inferences about the probability of an item being remembered in the future. Thus, although higher JOL accuracy is not ensured, the accuracy is achieved through the association between the cues utilized in forming JOLs and the performance on criterion tests (Dunlosky &

Metcalfe 2009). Regarding this, Koriat (1997) suggested that JOLs will be more affected by intrinsic cues (characteristic of an item) rather than extrinsic cues (including learning conditions).

In summary, metacognitive judgments and the factors affecting these judgments have been discussed above. When the factors affecting these judgments are considered, the question of how these judgments will change within the framework of possible memory errors comes to mind. Therefore, in the following sections memory fallibility and its possible relationship with metacognitive judgments will be discussed.

1.4 Memory Fallibility

Exploring memory involves acquiring knowledge, retaining information, forgetting, and memory errors. While memory is generally reliable, studies show it can be fallible. Hence, different approaches to memory errors and forgetting are found in literature. In the Seven Sins of Memory article, Schacter (1999) described forgetting as transience, absent-mindedness, blocking, misattribution, suggestibility, bias, and persistence. Transience, absent-mindedness, and bias relate to forgetting, while blocking, misattribution, and suggestibility refer to memory distortions. Time makes events less accessible (transience), attention affects forgetting (absent-mindedness), and information can become temporarily inaccessible (blocking). Misattribution occurs when the source of information is wrongly attributed. Suggestibility leads to false memory formation through suggestions. Bias is influenced by preexisting knowledge, and persistence involves remembering something unwanted. Schacter's explanations govern countless studies on memory errors. One of the phenomena of forgetting related to these theories is retrieval-induced forgetting (RIF). RIF refers to the process in which repeatedly recalling certain information from memory can lead to the deterioration of related long-term memories, resulting in forgetting (Anderson et al., 1994). In the original RIF experiment, participants first learned a set of category-cue pairs (e.g., fruit-banana, fruit-orange, color-blue). They then went through a retrieval practice (RP) phase, where they were presented with a category cue (e.g., fruit) and asked to recall a related item that they had studied (e.g., "What fruit did you

study beginning with 'or?"). Half of the category-cue pairs were practiced multiple times (Rp+), while the other half was not practiced at all (Rp-). After the RP phase, participants were asked to recall as many items as they could from all the categories they had studied. The main finding of this study was that participants had a worse recall for the unpracticed members of a given category compared to items from categories that were not practiced at all (Nrp). This suggests that the process of selecting certain items to remember also involves inhibiting the recollection of other items within the same category, leading to the decreased accessibility of those items (Anderson et al., 1994).

Furthermore, since the importance of relying on memory, especially when making critical decisions, is indisputable, many of the studies on forgetting and memory distortions in history have emerged specifically on eyewitness memory. Wells and Loftus (2003) indicated that an event involving being an eyewitness leaves a trace in the brain. Therefore, this trace can be destroyed, lost, or contaminated resulting in incorrect judgment about the questioned event (Wells & Loftus, 2013). For instance, Shaw et al. (1995) use an eyewitness memory test including colored slides about a crime scene in a student room to examine the inhibitory effect of retrieval on eyewitness memory. After the participants presented the slides, researchers rapidly questioned the participants, related to the object categories that were presented in the scene (e.g. clothes) which resulted in the retrieval of the target objects (Shaw et al., 1995). However, other object categories are not questioned by the experimenter (e.g. stationaries). Shaw et al. (1995) found that objects that are not retrieved or reviewed are recalled less than the objects from the repeatedly investigated. Concluding that successful retrieval of related objects blocked the accessibility to objects that were not rapidly reviewed.

The concept of misinformation effect occurs when post-event information distorts an individual's memory, changing reported details (Challies et al., 2011). Schacter (1999) found that misleading questions can alter memories of past events. In Loftus and Pickrell's (1995) experiment participants were given false information about a crime scene, leading them to adopt the misinformation as correct. Also, Wells and Bradfield (1998) demonstrated how misleading feedback to eyewitnesses can increase confidence in memories and result in more detailed recollections of events. Overall, research shows that post-event information and suggestions can significantly

impact the accuracy of memories and witness testimonies. Later studies examined the relationship between RIF and the misinformation effect. Many studies support the inhibition account of RIF suggesting that when cues trigger responses that are incompatible with the current context, it is thought that these responses are suppressed to enable the retrieval of the intended responses (Storm & Levy, 2012). After studies on the fact that less recall of items is due to inhibition, it has been a matter of debate whether this mechanism is also effective in the formation of the misinformation effect. Hence, later studies examined the relationship between RIF and the misinformation effect. Consequently, Macleod and Saunders (2005) clarified that when misinformation is presented about materials thought to be inhibited, the probability of recalling this misinformation is found to be higher than the probability of recalling the target information. Implying that their finding was compatible with the inhibitory account of RIF (Macleod & Saunders, 2005).

Further, it can be thought that examining the relationship between these memory fallibilities may reveal the sensitivity of metacognition in monitoring memory distortions. For instance, a reassessment of metacognitive judgments about one's memory can be made when a memory error such as forgetting or inaccurately recalling information is detected. Accordingly, adjustments in cognitive strategies can be made.

1.5 Metacognition and Memory Fallibility

Although studies have generally closely examined the relationship between memory fallibility and various cognitive functions, more studies must be needed to examine their relationship with metacognitive processes. However, existing studies examining the fallibility of memory in the context of metacognition also highly include eyewitness memory. Thus, Hollins and Weber (2016) indicated that eyewitness metacognition is studied broadly with event memory (concentrates metacognitive control) and identification decisions (focusing on metacognitive monitoring) in the literature. Also, it is highlighted that while event memory indicates recalled-based memory tests, identification indicates recognition tests (Hollins & Weber, 2016).

Dunlosky and Metcalfe (2009) argue that suggestive questions can lead people to believe false information as true, indicating a failure in metacognitive monitoring or decreasing metacognitive judgments. Metamemory monitoring involves individuals judging their memory's contents and source (Hart, 1967). Misinformation can confuse source monitoring, leading to memory errors (Metcalfe & Shimamura, 1994). Also, Wells and Loftus (2013) note that memory errors from suggestive information stem from issues in source monitoring. Loftus et al. (1989) found that misinformed individuals confidently report incorrect details as accurate memories. Dunlosky and Metcalfe (2009) suggest that repetition can increase memory confidence, regardless of accuracy. In sum, misleading post-event information can generate false memories that are confidently maintained. Furthermore, one study examining the RIF and JOL indicated that participants showed higher JOL scores for NRP items rather than RP-items (Carroll et al., 2007). Thus, Carroll et al. (2007) emphasize that although RIF is assumed to be an unconscious process their findings may provide evidence that the RIF process is monitored by metacognition. Also, this finding was compatible with the retrieval-fluency hypothesis suggesting that having easy access to more information about the desired response will lead to greater confidence in one's memory for it (Dunlosky & Metcalfe, 2009). On the other hand, Zhou et al. (2016) concluded that a decrease in inhibition function results in impaired metacognition by highlighting the significant correlation between the inhibition function and JOL accuracy. Since high activation in inhibition mechanisms is thought to be the reason behind the RIF effect, this finding may also be an example pointing out the possible relationship between the RIF and JOLs. For instance, it can be thought that if metacognitive monitoring is successfully maintained, it could be possible for individuals to have an accurate sense of learning material. Thus, higher JOL accuracy can be achieved by participants recognizing inhibited materials providing lower judgments for them, and providing higher judgments for practiced materials.

Although no study has determined the direct relationship between RIF and the FOK, as suggested by the target accessibility theory of FOK, accessing a larger quantity of information about a sought-after target, results in increased memory confidence in recognizing it (Dunlosky & Metcalfe, 2009). Also, Koriati (1993) suggested with the accessibility hypothesis that regardless of the accuracy of the retrieved information, the FOK judgments will be higher as individuals recall more.

Thus, it seems possible that there is a relationship between RIF and the FOK judgments as retrieval practice provides access to rapid information about the target item. Therefore, RIF can affect the FOK judgments by altering an individual's ability to judge their memory abilities and knowledge accurately. For instance, repeatedly retrieving information from memory may cause individuals to have a false sense of knowing better than they do. Thus, increasing confidence in their memories could lead to subsequent overestimation of their FOK. On the other hand, if the awareness of the individuals about the retrieval-induced phenomenon is increased, this may lead to a more accurate assessment of their memories. For instance, if an individual realizes that they cannot remember some of the items related to the practiced category, they can be more cautious when they assess their FOK. Since the underlying mechanism of RIF is effective in the misinformation effect, similar effects can be assumed on the metacognitive judgments. Additionally, the literature includes interventions to improve memory despite the fallibilities in memory (e.g. providing feedback about one's performance or providing memory cues). For instance, individuals' recognition of potential memory distortion can be understood from the changes in memory judgments, and in these cases, individuals can seek aids to enhance memory performance. Therefore, in the following section, the effects of cues and feedback will be discussed regarding memory distortions and metacognitive judgments.

1.6 Memory Performance with Cues and Feedback

Various studies focus on enhancing information recall, with retrieval cues playing a key role. Tulving and Osler (1968) highlighted the significance of retrieval cues in facilitating the recall process. Researchers have examined the impact of cues on RIF, with conflicting results. Perfect et al. (2004) found that using the same cues during testing and practice stages strengthened the forgetting effect. However, when fewer cues were used, no significant difference was observed. These results suggest that RIF is context-specific. Butler et al. (2001) also explored cueing techniques in the RIF paradigm, finding that the recall rate did not differ significantly based on cue availability for different types of items. Overall, the use of retrieval cues influences

the effectiveness of memory recall processes. Additionally, Koriat (1997) reported that participants do not assess memory strength when making JOL or FOK. Instead, cues such as beliefs about memory efficacy, encoding strategies, and study trials influence memory performance and memory prediction. External factors such as repeating lists have a weaker impact on JOL compared to recall, indicating these factors are not considered in forming JOL. Previous studies indicated that feedback could be beneficial for correcting memory errors (Butler et al., 2008; Sitzman et al., 2013). Feedback can also enhance memory retention, identify areas for further study, and highlight errors that need correction. It also provides diagnostic cues for future actions, helping learners quickly fix the errors (Sitzman et al., 2016). Furthermore, Ren et al. (2023) tested the feedback effect on prospective memory performance with three conditions (ongoing-test feedback, prospective memory feedback, and no feedback). Improvement in prospective memory performance was observed via provided feedback. Researchers also suggested that condition-related fluctuations in prospective memory performance indicate dynamic strategic monitoring (Ren et al., 2023). Also, Greenspan and Loftus (2020) reported that significant confidence change is observed after feedback about the performance. However, literature also includes contradictions about the subject. For instance, the provided feedback impairs the accuracy of participants' predictions about future test performance (Kornell & Rhodes, 2013).

1.7 The Current Study

Studies have shown that memory is less reliable than previously believed, with errors and missing information affecting our daily lives. Research focuses on RIF and misinformation, especially in eyewitness memory, essential for critical decisions. Memory accuracy can be improved by cues and feedback prompting individuals to reassess their decisions (Vortac et al., 1995; Zhang et al., 2016). Understanding these processes' impact on metacognitive judgments is crucial for developing strategies to enhance memory reliability. Further investigation is needed to explore new approaches to increasing memory accuracy.

The above-cited literature stresses the relationship between RIF, misinformation, memory cues, providing feedback, and metacognitive judgments. On the other hand, it is not clear how metacognitive judgments, including the misinformation effect, change in the RIF paradigm and how cues and feedback affect memory performance. Moreover, it needs to be clarified how metacognitive judgments, including the misinformation effect, change in the RIF paradigm and how cues and feedback affect memory performance. Although the task used in this study is not a typical eyewitness memory-related task, by expecting participants to remember the witnessed details of the presented text, it can provide valuable insight into memory processes of RIF, misinformation effect, and metacognitive judgments, which are relevant factors in the understanding of eyewitness testimony. Therefore, this study is expected to contribute to the literature by examining prospective and retrospective memory judgments and by working on manipulating misinformation under eyewitness testimony. This study will contribute to the theoretical approach, including contradicting assumptions about these phenomena.

The main objective of this study was to examine the metacognitive judgments concerning the actual memory performance in the effect of RIF and misinformation. To achieve this objective similar task used by MacLeod and Saunders (2005) related to RIF and misinformation task was implemented. During the study phase of this task, participants were presented with a brief background story involving robberies that took place in two different houses, followed by the items stolen from the house (in brief informational sentences). Afterward, half of the stolen items from a house (three different descriptions for each item) were presented to the participants as a retrieval practice. After this practice, the participants performed the cued or free recall phase. Following, participants were presented with misinformed sentences, and for the next phase, they were asked to choose the target from the four options.

However, item-based JOL judgments were implemented into the paradigm after the retrieval practice, and item-based FOK judgments were implemented into the paradigm after the misinformation phase. Moreover, this study aims to observe changes in individuals' overall confidence in their memories after retrieval practice and misinformation phases. To accomplish this goal, questions regarding eliciting participants' overall memory confidence were inserted before and after the retrieval practice and at the end of the study. Another aim of this study was to explore the

influence of cues and fallacious feedback on metacognitive judgments in the context of RIF and misinformation effects. To achieve this objective, before the task started, participants were randomly assigned to one of the four task conditions: cue/feedback present, cue/feedback absent, cue present /feedback absent, and cue absent/ feedback present. Thus, additional cues were inserted in the recall phase (by presenting the same sentences in the study phase with a blank instead of the target word), and fallacious feedback including questions related to participants' performance inserted after the forced choice recognition phase.

Hence, there are 3 main hypotheses are presented regarding recall, recognition, and metacognitive judgments in this present study.

First, the recall and recognition performances along with JOL and FOK judgments for NRP items are expected to be higher than the RP- items. Second, it is expected that the recall and recognition performances, JOL, and FOK judgments to be higher for participant groups that are presented with the cue. Lastly, while the overall memory judgments are expected to increase after the retrieval practice, it is also expected that the overall memory judgments of the participant group that provided fallacious feedback to be lower than the participants in other testing conditions.

Chapter 2

Method

2.1. Participants.

One hundred thirty-two healthy young adults (95 female and 36 male) who are Native Turkish speakers participated in this study voluntarily. The educational backgrounds of the participants ranged from holders of high school degrees to bachelor's and graduate degrees. The age range was 18-32, with a mean of 23.70 (SD=3.88). The snowballing sampling technique was used to recruit the participants. Following exclusion criteria were applied: Participants who reported a history of mental and neurological health problems and impaired vision scored above 17 in the Beck Depression Inventory and participants who scored above 21 in the Beck Anxiety Inventory. Participants' demographic information among the study conditions was presented in Table 1.

2.2. Materials

2.2.1 Form of sociodemographic information. This form was used to obtain the participant's sociodemographic information including, age, sex, occupation status, level of education, eye health, and status of previously experienced mental and neurological illnesses. If any mental and neurological illnesses were experienced it asked participants to report the type of the illness and, the name of the drugs if they were using any for its treatment. Further, they were asked to report whether they have eye-related problems or not and whether they are using any glasses or lenses for their treatment. Also, participants were asked to report whether the use of glasses eliminated their problems (See Appendix B).

2.2.2. Beck anxiety inventory. To evaluate the level of anxiety of the participants, the Beck Anxiety Inventory, is a self-report questionnaire consisting of 21 items for individuals aged 17 and above. It was developed by Aaron T. Beck and is used to assess anxiety levels. In Turkey, validity and reliability studies were conducted by Ulusoy and colleagues (1998) In this inventory responses to its items are scored between 0 and 3. Responses are evaluated as 0 points for "Not at all," 1 point for "Mild," 2 points for "Moderate," and 3 points for "Severe." Therefore, Higher scores on this inventory indicate higher levels of anxiety (See Appendix C).

2.2.3 Beck depression inventory. This self-report questionnaire consisting of 21 items was created by Beck et al. in 1961 and is used to assess depression levels. Participants were requested to assess items based on their feelings over the past week, including the day of responding to the questionnaire. While the minimum score for each question was 0, the maximum score was 3. Thus, the total score of BDI ranges from a minimum of 0 to a maximum of 63. Higher scores indicate higher levels of depression. The standardization of this scale for Turkish was made by Hisli (1988) to assess the participants' depression levels. A score of 17 and above on this scale indicated clinical depression (Hisli, 1988) (See Appendix D)

2.2.4 The retrieval-induced forgetting and misinformation task. This task is a revised version of a task used by MacLeod and Saunders (2005). It is also a computerized memory task that incorporates evaluations of metacognitive judgment within the same subtasks and guidelines as the original task. In addition, this task includes cues presented to the participants during the recall phase and fallacious performance feedback provided at the end of the study (see Figure 1).

Thus, this task consisted of study, retrieval practice, recall, misinformation, forced-choice recognition, and feedback phases with metacognitive judgment evaluations. The test took an average of 45 minutes, depending on the progress performance of all participants. The duration required to read and answer the study materials was determined based on the average data of 15 individuals who participated in the preliminary study and reported the time taken to read and respond to the materials. Additionally, after the study was translated into Turkish and some additions were made to align with the nature of the study, these translations were evaluated by

three experts in the field. They were subsequently revised based on the received feedback.

2.2.5 Study phase. This phase included a background story, target items, and retrieval practice. MacLeod and Saunders (2005) used a similar backstory about burglary in this phase. Therefore, two background stories about a burglary from different households (Özkan's house and Yılmaz's house) and 40 items (20 items per house) with their brief descriptions of what had been stolen were given to the participants (See Appendix E). For the retrieval practice, participants were presented with three different descriptions of half of the items stolen from Özkan's house (30 descriptions) (See Appendix F).

2.2.6 Recall phase. For this phase, some participants were presented with empty boxes for each of the stolen items (40 items) and asked to do free recall. In comparison, others were presented with the original sentences with the target word blank as a cued recall (40 items).

2.2.7 Misinformation task. Misleading questions about the target items were presented to the participants. A total of 30 misinformation sentences were presented to the participants about the practiced items from the practiced household, unpracticed items about the practiced household, and items from the unpracticed household (See Appendix G). Participants were asked to answer misleading questions (e.g. Where in the house was the *plant* stolen? Target was the *vase*).

2.2.8 Forced choice recognition task. Questions about the target items were presented to the participants. The same sentences in the study phase were provided blank in place of the target item with four different options, and participants were asked to select the correct answer from the four choices.

2.2.9 Metacognitive judgment rating. JOL appraisals was made based on the given answers on the recall phase item by item. Also, the FOK judgment was made based on future answers in the recognition phase. Participants' judgments were obtained through a six-point Likert scale.

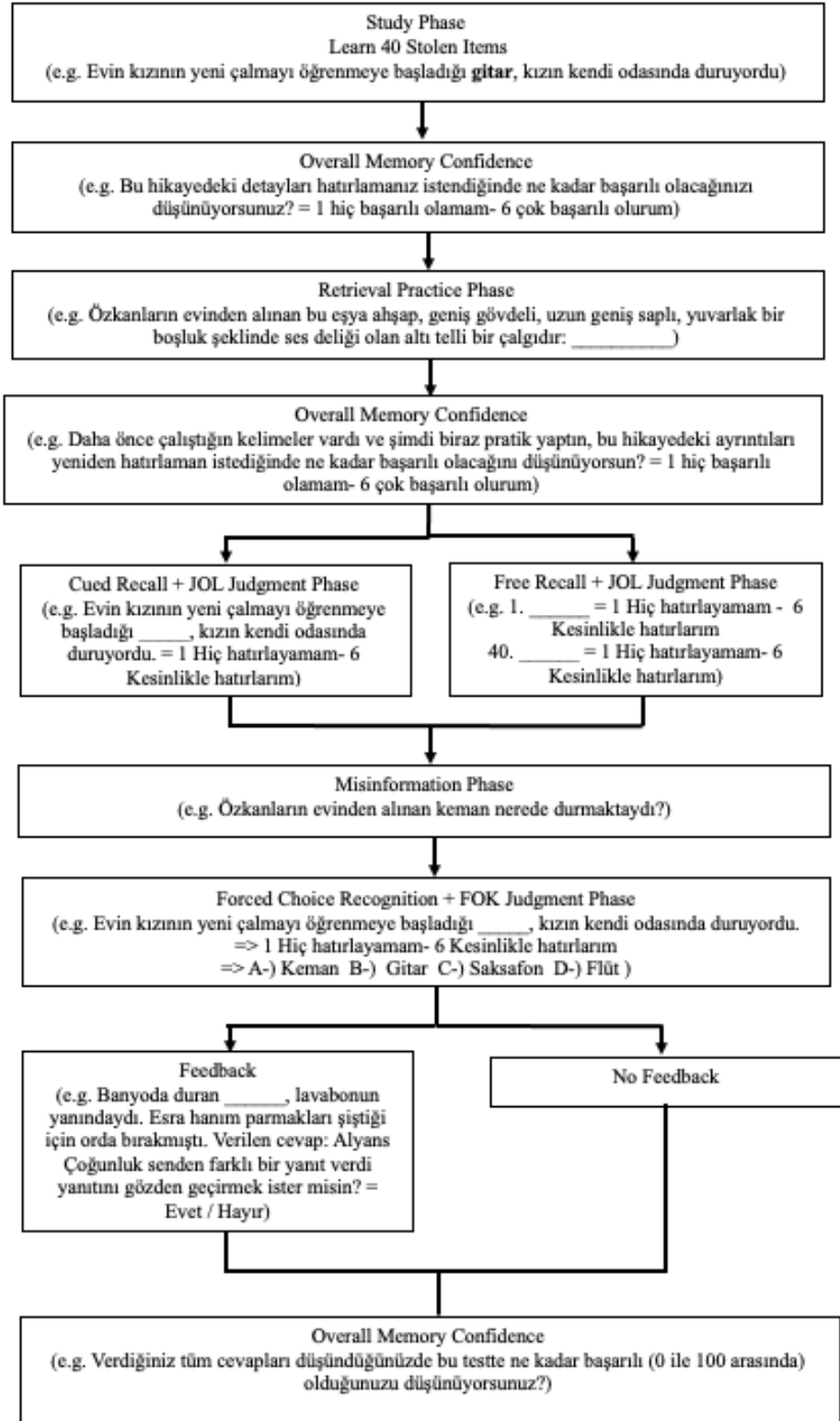


Figure 1. A Schematic Depiction of the Task Was Used in the Study

2.2.10 Rating of subjective confidence judgment. The overall retrospective memory judgment indicating participants' confidence about their performance was made before and after the retrieval practice on a 6-point Likert scale (1 for 'I am not sure at all, 6 for 'I am completely sure'). Also, the overall prospective memory judgment indicating participants' confidence in the whole test achievement was made at the end of the test on a percentage scale.

2.3 Procedure

Firstly, informed consent was given to the participants (See Appendix A), which provided information about the study (possible risks, benefits, and content). After the informed consent, a sociodemographic information form was given along with the Beck anxiety and Beck depression inventories. Since this study was designed to be conducted online via computer, individuals participated online via the created website.

Later, participants were randomly assigned one of the four conditions: given cue, no cue, feedback, or no feedback. After that participants began to read a study about the burglary. In the study phase participants were asked to read two short stories related to a burglary in two different households. Participants first read a paragraph that included what the family was doing on the day of the robbery and environmental evidence of the robbery. To read the one-paragraph-long background information about the day of the robbery, 30 seconds were given to Özkan's household, and 32 seconds were given to the Yılmaz's household. Following the informative paragraph about the robbery day, a list of 20 items (per household) that had been stolen was presented to the participants one by one with brief descriptions. A brief description included details like where that item was placed or who owned the item (see Appendix E). Also, every sentence stayed on the screen for a maximum of 10 seconds depending on the length of the sentence. The overall memory judgment question appeared after the presentation of the background story and the list of stolen items. Participants answered the question, "How successful do you think you will be when asked to recall details from this story?" on a six-point Likert scale.

After that, for retrieval practice, participants practiced half of the stolen items three times from one of the households. The three descriptions of the same item appeared on the screen, and participants answered by writing that target item (see Appendix F). After the retrieval practice phase, another overall memory judgment question appeared on the screen. Participants answered the question, "You had words that you studied before, and now you have practiced a bit, how well do you think you will be successful if you're asked to recall details from this story?" on a six-point Likert scale. After this, a distractor task was presented to the participants. They were given three minutes to generate 1 animal name corresponding to each letter of the alphabet.

After the presentation of the first distractor, the task continued with the recall phase. In this phase, some participants were asked to write all the items they remembered. At the same time, participants in the cued recall condition presented the same sentences in the study phase with a blank in place of the target item. They asked to fill in the blank with the item's name that they could remember. Following the recall phase of every item, an item-based Judgement of learning scores was taken from the participants. Participants were asked to assess the likelihood of correctly recalling each item on a six-point Likert scale. Later second distractor task was presented. Three minutes were given to the participants to write Blue, black, rounded, wooden objects corresponding to every letter in the alphabet. After that in the misinformation phase, the participants were exposed to questions containing misinformation about the practiced items from the practiced households (items that were presented in the retrieval practice phase from Özkan's House), non-practiced items from the practiced household (Other items in Özkan's house) and not the practiced household (Yılmaz's house) (see Appendix G). Later as a final distractor participant was asked to write all the country names corresponding to every letter in the alphabet. Three minutes were given to the participants for this phase.

After that in the FOK judgment and forced choice recognition phase, the sentences shown to the participants about the stolen items during the study phase were presented again with the blank space instead of a target word. Participants answered the "Can you remember if the correct answer for this part in the blank was given to you in the future in a multiple-choice option?" question on a 6-point Likert scale to obtain their FOK judgments for every item. After the presentation of every sentence "Which item should be placed in the blank?" question appeared and participants were

asked to select the correct answer between four choices. These four choices included one target item, one misinformed item, and two lured items. Later, After the forced choice recognition phase is over, some participants are presented with 10 fallacious feedback depending on the assigned conditions. Thus, some participants answered yes or no to the question "Others gave a different answer than you. Do you want to review your answer" that they saw for 10 of the answered forced-choice questions. Other participants did not receive any feedback about their performance. After that, the last overall memory judgment question appeared on the screen and participants answered out of 100 to the question, "When you think the answers that you give, how many of the answers do you think you remember correctly?"



Chapter 3

Results

Before beginning the statistical analysis, data screening procedure was conducted. The mean variables (total correct for recall and total correct for recognition) skewness and kurtosis values were checked and both values were in the accepted range. To determine whether data from participants who scored above 17 on the BDI and above 21 on the BAI should be excluded, both sets of data (with scores exceeding the cutoff point and those without) were compared in terms of their means. Although no main difference was detected between the two sets of data, to be able to prevent possible confounding variables or outliers associated with higher BDI and BAI scores, participants with BDI and BAI scores exceeding the cutoff point were excluded. According to the Mahalanobis distance with a criterion of $\alpha = .01$, $\chi^2 = 6.63$, six outliers among the cases were found and removed from the data. After data screening, a total of 59 participants were removed from the data. Thus, data from the remaining 132 participants were analyzed. Those participants' demographic information is presented in Table 1.

3.1 Results for Recall Performance

We first hypothesized that recall performances for NRP items would be better than the RP- items implying an RIF effect for the recall phase and recall performance of the participants in the cued conditions would be better than the participants in the non-cued conditions. Thus, to test the effect of retrieval-induced forgetting in the recall phase, as a manipulation check, correct answers from three item categories were compared with repeated measures of ANOVA. In the analysis, the dependent variable was the number of correctly answered questions while the independent variable was item categories. The results showed that (see Table 2) there was a significant difference between the number of correctly answered questions of RP+, RP-, and NRP items $F(2,248) = 91.86$, $p < .001$. Bonferroni test was used for posthoc comparisons. It was found that the three groups exhibited significant

differences from each other. A significantly higher recall performance is observed for items from the RP+ category ($M = 5.90$) compared to items from the RP- ($M = 2.13$) and NRP ($M = 4.57$) categories.

Table 1

Demographics of the Participants According to Study Conditions

Variables	Cue/ Feedback Present		Cue/Feedback Absent		Cue Present Feedback Absent		Cue Absent Feedback Present	
	<i>n</i>	<i>M (SD)</i>	<i>n</i>	<i>M (SD)</i>	<i>n</i>	<i>M (SD)</i>	<i>n</i>	<i>M (SD)</i>
Gender								
Female	24		25		26		20	
Male	7		11		9		9	
Age	31	23.45(2.73)	37	23.59(4.60)	35	24.60(2.54)	29	23.03(5.10)
Education Level								
High School	9		9		8		8	
Graduate								
Bachelor	17		22		27		16	
Graduate								
Masters	5		1		5		5	
Graduate								

Another repeated measure ANOVA is conducted to explore differences between incorrectly recalled questions from three different item categories to support the findings for the first hypothesis. Thus, the dependent variable was the number of incorrectly recalled questions while the independent variable was item categories. A significant difference was found $F(2,258) = 11.39$, $p < .001$. According to the Bonferroni test that is conducted for posthoc comparisons, better recall performance was observed for items from the RP+ category ($M = 1.03$) compared to items from the RP- category ($M = 1.75$) and NRP category ($M = 2.16$). No other group differences were observed. See Table 2 for means and standard errors for correctly and incorrectly recalled questions based on item categories.

Furthermore, a one-way ANOVA is carried out to test the effects of testing conditions on the total number of correctly answered questions (including all questions that were correctly answered regardless of their categories). Thus, while the independent variable was the testing conditions (cue/ feedback present, cue/ feedback absent, cue present/feedback absent, cue absent/ feedback present) the dependent variable was the total number of correctly answered questions. The effect of the testing condition was not significant $F(3,126) = .025, p > .05$.

Table 2

Difference of Item Categories on Recall. Repeated measure ANOVA Results with descriptive statistics

Variables	RP+	RP-	NRP	F
	M (SE)	M (SE)	M (SE)	
Correct Recall	5.90 (.21)	2.13(.18)	4.57 (.32)	91.86**
Incorrect Recall	1.03 (.16)	1.75 (.22)	2.16 (.32)	11.39**

Note. RP+: practiced items from the practiced category; RP-: non-practiced items from the practiced category; NRP: non-practiced items from the non-practiced category.

** $p < .01$, * $p < .05$

Another one-way ANOVA is conducted to see whether there is an effect of testing conditions on the total number of incorrectly answered questions. While the independent variable was testing conditions, the dependent variable was the number of incorrectly answered questions. Results indicated significant differences between the four testing conditions $F(3,126) = 31.43, p < .001$. According to posthoc comparisons, participants' recall performance in the cue absent / feedback present testing condition ($M = .82$) was better than the participants in the cue/ feedback present ($M = 7.44$) and cue present/feedback absent condition ($M = 11.14$). Also, participants' recall performance in the cue/ feedback absent condition ($M = 1.37$) was better than the participants in the cue/ feedback present ($M = 7.44$) and cue present/ feedback absent condition ($M = 11.14$). See Table 3 for means and standard errors for correctly and incorrectly recalled questions based on testing groups.

Lastly, the amount of inhibition that each participant demonstrated was calculated by subtracting correctly answered RP- items from NRP items as a

measurement of RIF. Therefore, one-way ANOVA is conducted to see whether the degree of RIF in recall performance differs according to the four testing conditions. The results indicated no significant difference between testing conditions in terms of RIF degree in recall performance $F(3,128) = .763, p > .05$ (See Table 3).

Table 3

Effects of Testing Condition on Recall Performance. One-Way ANOVA Results with Descriptive Statistics

Variables	Cue/ Feedback Present	Cue/Feedback Absent	Cue Present Feedback Absent	Cue Absent Feedback Present	<i>F</i>
	<i>M (SE)</i>	<i>M (SE)</i>	<i>M (SE)</i>	<i>M (SE)</i>	
Correct Recall	11.72 (1.04)	11.62 (1.17)	12.02(1.07)	11.75(1.20)	.025
Incorrect Recall	7.44 (1.01)	1.37 (.40)	11.14 (1.34)	.82(.36)	31.43*
Amount of Inhibition	1.80 (.50)	2.62 (.55)	2.82 (.54)	2.79 (.54)	.763

** $p < .01$, * $p < .05$

3.2. Results for Recognition Performance

The second hypothesis of the study is that recognition performances for NRP items would be better than the RP- items implying continuity of the RIF effect for the recognition phase and the recognition performance of the participants in the cued conditions would be better than the participants in in the non-cued conditions. Thus, to test the RIF effect in the recognition phase, as a manipulation check, correct answers given to item categories were compared with repeated measures of ANOVA. In the analysis, the dependent variable was the number of correctly answered questions while the independent variable was the item categories. The results showed that there was a

significant difference between the number of correctly answered questions of RP+, RP-, and NRP items $F(2,246) = 59.49$, $p < .001$. When applying a Bonferroni test for posthoc comparisons, it was found that the three groups exhibited significant differences from each other based on related t-tests. Significantly better recognition performance is observed for items from the RP+ category ($M = 6.99$) compared to items from the NRP ($M = 6.35$) and RP- ($M = 4.62$) categories. Results confirmed the continuing RIF effect after the misinformation phase.

Another repeated measure ANOVA is conducted to explore whether there are differences between item categories in terms of incorrect recall to support the previous findings. In the analysis, the dependent variable was the number of incorrectly answered questions while the independent variable was the item categories. A significant difference was found in $F(2,244) = 56.85$, $p < .001$. It was found that all three groups exhibited significant differences. The worst recognition performance was observed for items from the RP- category ($M = 4.60$) compared to items from the NRP ($M = 3.13$) and RP+ ($M = 2.29$) categories. See Table 4 for means and standard errors for correctly and incorrectly recognized questions based on item categories.

Furthermore, a one-way ANOVA is carried out to test the effects of testing conditions on the total number of correctly answered questions (including all questions that were correctly answered regardless of their categories) for the recognition phase. Thus, while the independent variable was the testing conditions (cue/ feedback present, cue/ feedback absent, cue present/feedback absent, cue absent/ feedback present) the dependent variable was the total number of correctly answered questions. A non-significant group difference was found between the four groups in terms of correctly answered questions in the recognition phase $F(3,120) = .750$, $p > .05$.

Table 4

Difference of Item Categories on Recognition. Repeated measure ANOVA Results with descriptive statistics

Variables	RP+	RP-	NRP	F
	M (SE)	M (SE)	M (SE)	
Correct Recognition	6.99 (.17)	4.62 (.20)	6.35 (.18)	59.49**
Incorrect Recognition	2.29 (.14)	4.60 (.21)	3.13 (.16)	56.85**

Note. RP+: practiced items from the practiced category, RP-: non-practiced items from the practiced category; NRP: non-practiced items from the non-practiced category.

** $p < .01$, * $p < .05$

Another one-way ANOVA is conducted to see whether there is an effect of testing conditions on the total number of incorrectly answered questions in the recognition phase to support the previous findings. In this analysis while the independent variable was testing conditions, the dependent variable was the number of incorrectly answered questions. Results indicated differences between the four test groups was non-significant $F(3,120) = .198, p > .05$ (See Table 5).

Lastly, to see whether the degree of RIF in recognition performance differs according to the testing conditions one-way ANOVA is conducted. In this analysis while the independent variable was testing conditions, the dependent variable was the degree of inhibition (calculated by subtracting correct answers to RP- from NRP items). The results indicated no significant difference between testing conditions in terms of RIF degree in recognition performance $F(3,128) = .422, p > .05$. See Table 5 for means and standard errors for the degree of inhibition in recognition based on testing groups.

Table 5

Effects of Testing Condition on Recognition Performance. One-Way ANOVA Results with Descriptive Statistics

Variables	Cue/ Feedback Present	Cue/Feedback Absent	Cue Present Feedback Absent	Cue Absent Feedback Present	<i>F</i>
	<i>M (SE)</i>	<i>M (SE)</i>	<i>M (SE)</i>	<i>M (SE)</i>	
Correct Recognition	18.16 (.94)	17.66 (.84)	17.84 (.72)	16.46 (.79)	.750
Incorrect Recognition	10.93 (.82)	9.30 (.72)	11.15 (.63)	9.28(.71)	1.98
Amount of Inhibition	2.03(.54)	1.48(.44)	1.68(.45)	1.31(.44)	422

** $p < .01$, * $p < .05$

3.3. Results for Metacognitive Judgments

3.3.1. Results for JOL judgments. It is also hypothesized that JOL and FOK judgments for NRP items will be higher than the RP- items and JOL and FOK judgments of participants in the cued conditions would be higher than the participants in the non-cued conditions. Thus, to test this hypothesis firstly a repeated measures ANOVA is carried out to observe whether JOL judgments for correct answers differ based on the question categories. In this analysis, the independent variable was the item categories while the dependent variable was the participants' JOL judgments for correctly answered questions. Results indicated non-significant differences between question categories regarding JOL judgments for correct answers $F(2,182) = 1.35, p > .05$ (See table 6).

To support the previous results, another repeated measure ANOVA is conducted to test whether JOL judgments for incorrect answers differ based on the question categories. Thus, in this analysis, the independent variable was the item categories while the dependent variable was the participants' JOL judgments for

incorrectly answered questions. Results indicated non-significant differences between question categories in terms of JOL judgments for incorrect answers $F(2,76) = .429, p > .05$. See Table 6 for means and standard errors of JOL judgments of correct and incorrect answers on a question category basis.

Table 6

Difference of Item Categories on JOL Judgments. Repeated measure ANOVA Results with descriptive statistics

Variables	RP+	RP-	NRP	F
	M (SE)	M (SE)	M (SE)	
JOL Judgments for Correct Answers	4.54 (.14)	4.43 (.15)	4.65 (.15)	1.35
JOL Judgments for Incorrect Answers	2.70 (.22)	2.85 (.17)	2.81 (.20)	.429

Note. RP+: practiced items from the practiced category, RP-: non-practiced items from the practiced category; NRP: non-practiced items from the non-practiced category.

** $p < .01$, * $p < .05$

Moreover, it was hypothesized that participants' JOL judgments in the cued test condition would be higher than the participant's JOL judgments in the non-cued condition. Thus, one-way ANOVA is carried out to test the effects of testing conditions on participants' JOL judgments for total correct answers (regardless of the category of the items). Results indicate a significant effect of testing conditions on participants' JOL judgments for correct answers $F(3,122) = 10.13, p < .001$ (See table 7). Post-hoc comparisons indicated participants' JOL judgments in the cue/ feedback absent condition ($M = 5.06$) were higher than the participants in the cue/feedback present ($M = 4.13$) condition and cue present/feedback absent condition ($M = 3.77$). Also, the JOL judgments of participants in the cue/feedback present condition were higher than the participants in the cue present/ feedback absent condition. Another significant difference was that participants' JOL judgments in the cue absent/ feedback

present ($M = 5.20$) condition was higher than the participants in the cue/feedback present ($M = 4.13$) and cue present/feedback absent ($M = 3.77$) condition. Furthermore, the JOL judgments of participants in the cue/ feedback present condition were higher than the participants in the cue present/feedback absent condition.

Table 7

Effects of Testing Condition on JOL Judgments. One-Way ANOVA Results with Descriptive Statistics

Variables	Cue/ Feedback Present	Cue/Feedback Absent	Cue Present Feedback Absent	Cue Absent Feedback Present	
	<i>M (SE)</i>	<i>M (SE)</i>	<i>M (SE)</i>	<i>M (SE)</i>	<i>F</i>
JOL Judgments for Correct Answers	4.13 (.25)	5.06 (.22)	3.77 (.20)	5.20 (.21)	10.13**
JOL Judgments for Incorrect Answers	2.91 (.28)	4.11 (.44)	2.71 (.16)	3.32 (.55)	4.07**

** $p < .01$, * $p < .05$

To support the previous results, another one-way ANOVA is carried out to observe whether the JOL judgments of participants for incorrect answers differ depending on the testing groups. In this analysis, the independent variable was testing conditions, and the dependent variable was the JOL judgments for incorrect answers. Results indicate a significant effect of testing conditions on JOL judgments for incorrect answers $F(3,85) = 4.07$, $p < .05$. Post-hoc comparisons indicated that participants' JOL judgments for incorrect answers were higher in the cue/feedback absent condition ($M = 4.11$) than the participants in the cue/feedback present condition ($M = 2.91$) and participants in the cue present/feedback absent condition ($M = 2.71$). No other significant group differences were observed. (See Table 7) for means and standard errors of JOL judgments of correct and incorrect answers on a group basis.

Participants' FOK and JOL accuracy, which refers to how well people can predict their future memory performance, was calculated by using the Goodman–Kruskal's Gamma correlation, a commonly used measure of FOK and JOL accuracy (Nelson, 1984). For each participant, a Gamma correlation between the metacognitive judgments (between 1-6) and the response accuracy (correct or incorrect), was calculated. To investigate whether JOL accuracy scores differ according to the testing conditions, one-way ANOVA is carried out. The results indicate the non-significant difference between testing conditions regarding JOL accuracy score, $F(3,74) = 1.60$, $p > .05$. Also, to test whether regarding JOL accuracy score for RP+, RP-, and NRP items differ according to the testing conditions separate one-way ANOVAs are carried out. However, all comparisons were insignificant. In details, ANOVA results for RP+ items ($F(1,41) = 2.23$, $p > .05$), RP- items ($F(1,44) = .03$, $p > .05$) and NRP items ($F(1,41) = .197$, $p > .05$) were non-significant. On the other hand, Gamma correlations were not calculated due to the small number of answers (responses) in the RP+, RP- and NRP categories. See Table 8 for means and standard errors for Gamma Correlations on a group basis.

Lastly, a paired sample t-test was carried out between the JOL judgments that were given to correct and incorrect answers to observe whether participants can recognize correct and incorrect answers at some level and evaluate their learning judgments accordingly. Thus, higher JOL judgments for correct answers compared to incorrect ones were expected. A paired sample t-test indicated that participants JOL scores for correctly answered questions ($M = 4.29$, $SE = .14$) was higher than the incorrect answers ($M = 3.04$, $SE = .16$) with a mean difference of $M = 1.23$ ($SE = .13$), $t(86) = 9.87$, $p < .001$. Also, tested variables were significantly correlated ($r = .64$, $p < .001$)

Table 8

Effects of Testing Condition on JOL Accuracy. One-Way ANOVA Results with Descriptive Statistics

Variables	Cue/ Feedback Present <i>M (SE)</i>	Cue/Feedback Absent <i>M (SE)</i>	Cue Present Feedback Absent <i>M (SE)</i>	Cue Absent Feedback Present <i>M (SE)</i>	<i>F</i>
Gamma Correlations for Total	.54 (.10)	.66 (.17)	.41(.08)	.82 (.06)	1.60
Gamma Correlations for RP+ items	.71 (.12)		.41 (.13)		2.23
Gamma Correlations for RP- items	.65(.15)		.68 (.07)		.030
Gamma Correlations for NRP items	.64(.13)		.71 (.07)		.197

** p<.01 , *p<.05

3.3.2. Results for FOK judgments. To test our hypothesis indicating that participants' FOK judgments will be higher for NRP items a repeated measures ANOVA is conducted to observe whether FOK judgments for correct answers differ based on the question categories. Therefore, while the independent variable in this analysis is the item categories, the dependent variable is the FOK judgments for correctly answered questions. Results indicated significant differences between question categories in terms of FOK judgments for correct answers $F(2,240) = 4.94, p < .05$. The pairwise comparisons indicated significant differences. Thus, the FOK judgments for RP+ ($M = 3.76$) items were higher than the FOK judgments for- RP- ($M = 3.52$) items. Also, FOK judgments for, NRP ($M = 3.74$) items were higher than

the FOK judgments for RP- items. No other significant differences were observed in the related t-tests (See Table 9).

Additionally, similar repeated measure ANOVA is conducted for incorrect answers. Thus, while the independent variable for this analysis was item categories, the dependent variable was the FOK judgments for incorrect answers. Results indicated non-significant differences between question categories in terms of FOK judgments for incorrect answers $F(2,198) = 1.74, p > .05$. (See Table 9).

Table 9

Difference of Item Categories on FOK Judgments. Repeated measure ANOVA Results with descriptive statistics

Variables	RP+	RP-	NRP	F
	<i>M (SE)</i>	<i>M (SE)</i>	<i>M (SE)</i>	
FOK Judgments for Correct Answers	3.76 (.12)	3.52 (.13)	3.74 (.14)	4.93**
FOK Judgments for Incorrect Answers	3.24 (.16)	3.22 (.14)	3.03 (.15)	1.74

Note. RP+: practiced items from the practiced category, RP-: non-practiced items from the practiced category; NRP: non-practiced items from non-practiced category.

** $p < .01$, * $p < .05$

Furthermore, to test the part of the previously mentioned hypothesis indicating that participants in the cued conditions FOK judgments will be higher than the participants in the non-cued testing condition, one-way ANOVA is carried out to observe whether the FOK judgments of participants for correct answers differ depending on the testing groups. Thus, the independent variable in this analysis was testing conditions while the dependent variable was FOK judgments for total correct answers (regardless of their question category). Results indicate a non-significant group difference in FOK judgments of correct answers $F(3,120) = .766, p > .05$. (See table 10).

Subsequently, another one-way ANOVA is carried out to support the previous findings by observing whether the FOK judgments of participants for incorrect answers differ depending on the testing groups. Thus, while the dependent variable stays the same for this analysis, the independent variable changes as FOK judgments for incorrect answers. Results indicate a non-significant group difference in FOK judgments of incorrect answers $F(3,120) = 1.82, p > .05$. See Table 10 for means and standard errors of FOK judgments of correct and incorrect answers on a group basis.

Table 10

Effects of Testing Condition on FOK Judgments. One-Way ANOVA Results with Descriptive Statistics

Variables	Cue/ Feedback Present	Cue/Feedback Absent	Cue Present Feedback Absent	Cue Absent Feedback Present	<i>F</i>
	<i>M (SE)</i>	<i>M (SE)</i>	<i>M (SE)</i>	<i>M (SE)</i>	
FOK Judgments for Correct Answers	3.76 (.26)	3.90 (.22).	3.48 (.22)	3.50 (.24)	.767
FOK Judgments for Incorrect Answers	2.38 (.29)	3.62 (.21)	2.90 (.22)	3.07 (.25)	1.82

** $p < .01$, * $p < .05$

For the FOK accuracy score, one-way ANOVA results showed a significant difference between testing conditions, $F(3,118) = 5.06, p < .05$. According to posthoc comparisons, the FOK accuracy of the participants in the cue present/ feedback absent condition ($M = .40$) was better than the FOK accuracy of the participants in the cue/ feedback present ($M = .10$) and cue/feedback absent ($M = .08$) condition. No other differences were observed between the groups. Furthermore, to test whether the FOK accuracy score of RP+, RP-, and NRP items differ according to the testing conditions separate one-way ANOVAs are carried out. Results for RP+ items ($F(3.96) = 1.96, p$

> .05.), RP- items ($F(1,106) = .368, p > .05.$) and NRP items ($F(1,95) = 1.98, p > .05$) indicate a non-significant difference (See Table 11).

Table 11

Effects of Testing Condition on FOK Accuracy. One-Way ANOVA Results with Descriptive Statistics

Variables	Cue/ Feedback	Cue/Feedback Absent	Cue Present Feedback Absent	Cue Absent Feedback Present	<i>F</i>
	Present				
	<i>M (SE)</i>	<i>M (SE)</i>	<i>M (SE)</i>	<i>M (SE)</i>	
Gamma Correlations for Total	.10 (.11)	.08 (.08)	.42 (.06)	.29 (.07)	4.06*
Gamma Correlations for RP+ items	.01 (.18)	.11 (.13)	.45 (.11)	.19 (.13)	1.96
Gamma Correlations for RP- items	.05 (.13)	.04 (.12)	.20 (.12)	.07 (.14)	.360
Gamma Correlations for NRP items	.25 (.13)	.16 (.14)	.34 (.12)	.56 (.08)	1.98

** $p < .01$, * $p < .05$

Additionally, a paired sample t-test was carried out between the FOK judgments that were given to correct and incorrect answers to observe whether participants can recognize correct and incorrect answers at some level and evaluate their FOK judgments accordingly. Therefore, higher FOK judgments for correct answers compared to incorrect ones were expected. A paired sample t-test indicated that participants' FOK scores for correctly answered questions ($M = 3.66, SE = .12$) were higher than the incorrect answers ($M = 3.25, SE = .12$) with a mean difference of $M =$

.41 ($SE = .06$), $t(123) = 6.58$, $p < .001$. Also, tested variables were significantly correlated ($r = .86$, $p < .001$).

3.3.3 Results for overall memory confidence. The last hypothesis was expecting that participants' second overall memory judgments to be higher than the first overall memory judgments that are taken before the retrieval practice phase. Also, it was expected that the overall memory judgments of the participant group that provided fallacious feedback to be lower than the participants in other testing conditions. Thus, firstly, a paired sample t-test was carried out to observe whether retrospective memory judgment scores taken before and after the retrieval practice differed. The results showed that participants showed better retrospective memory judgment after the retrieval practice ($M = 2.43$, $SE = .12$) than before ($M = 2.90$, $SE = .13$). Hence, results indicated that retrospective memory judgment scores increased after the retrieval practice phase of the study with the main difference of $M = .46$, ($SE = .14$), $t(131) = 3.33$, $p \leq .001$ (See Table 12).

Table 12

Mean Comparison of Retrospective Memory Judgments After and Before Retrieval Practice

Variable	After		Before		$t(131)$	p	r
	M	SE	M	SE			
Retrospective							
Memory Judgment	2.90	.13	2.43	.12	3.33	.001	.36**

** $p < .01$, * $p < .05$

Subsequently, one-way ANOVA was conducted to see whether overall retrospective memory judgments and prospective memory judgments differ depending on the testing conditions. While the independent variable that remained the same across separate analyses was the testing conditions, the dependent variables were respectively the first retrospective memory judgment, the second retrospective memory judgment, and the prospective memory judgment. Hence, the results indicate

a non-significant difference between the first retrospective memory judgment ($F(3,128)=.118$, $p >.05$.), and second retrospective memory judgment ($F(3,128) = .527$, $p >.05$.) and prospective memory judgments of participants ($F(3,128) = 1.41$, $p >.05$.), in different testing groups (See Table 13).

Table 13

Effects of Testing Condition on Overall Metacognitive Judgments. One-Way ANOVA Results with Descriptive Statistics

Variables	Cue/ Feedback	Cue/Feedback Absent	Cue Present Feedback Absent	Cue Absent Feedback Present	<i>F</i>
	Present				
	<i>M (SE)</i>	<i>M (SE)</i>	<i>M (SE)</i>	<i>M (SE)</i>	
Retrospective Memory Judgment 1	2.51 (.26)	2.37 (.24)	2.48 (.20)	2.34 (.24)	.118
Retrospective Memory Judgment 2	3.12 (.27)	3.00 (.26)	2.74 (.23)	2.72 (.32)	.527
Prospective Memory Judgment	53.29 (.26)	43.54 (3.82)	44.71 (3.55)	42.75 (4.52)	1.41

** $p <.01$, * $p <.05$

3.4. Correlations Among Metacognitive Judgments

Although it was not among our main hypotheses, we tested correlations between metacognitive judgments. A Pearson correlation analysis was conducted between the JOL judgments, FOK judgment scores, first and second retrospective memory judgments, and prospective memory judgment variables. The correlation analysis revealed a significantly positive correlation between these variables. In detail, a moderate positive correlation was found between the JOL judgments and FOK judgments $r(117) = .524$, $p <.001$. Also, a relatively low positive correlation was found between the JOL judgments and the first retrospective memory

judgment $r(124) = .202, p < .05$, second retrospective memory judgment $r(124) = .317, p < .001$, prospective memory judgment $r(124) = .383, p < .001$. Furthermore, while a moderate positive correlation was found between the FOK judgment and second retrospective memory judgment $r(122) = .426, p < .001$, prospective memory judgment $r(122) = .635, p < .001$, a significantly low positive correlation was found between the FOK judgments and first retrospective memory judgment $r(122) = .268, p < .01$. The correlation analysis also revealed a significantly low positive correlation between the first retrospective memory judgment and the second retrospective memory judgment $r(130) = .366, p < .001$, prospective memory judgment $r(130) = .271, p < .01$. Lastly, a significantly low correlation is found between the second retrospective memory judgment and prospective memory judgment $r(130) = .406, p < .001$. See Table 14 for correlations between the metacognitive judgments.

Table 14

Correlations Among Metacognitive Judgments

Variable	(1)	(2)	(3)	(4)	(5)
(1) JOL Judgment	1	.524**	.202*	.317**	.383**
(2) FOK Judgment		1	.268**	.426**	.635**
(3) Retrospective Memory Judgment 1			1	.366**	.271**
(4) Retrospective Memory Judgment 2				1	.406**
(5) Prospective Memory Judgment					1

** $p < .01$, * $p < .05$

Chapter 4

Discussion

4.1. Summary of the Results

This present study aimed to explore the metacognitive judgments under the misinformation effect in the RIF paradigm with provided cues and feedback. To test our hypotheses, participants were assigned one of the four conditions (cue/feedback present, cue/ feedback absent, cue present/feedback absent, cue absent/ feedback present). Later, a RIF paradigm involving the misinformation phase was administrated. Following the study phase, after the retrieval practice, and at the end of the test phase , participants reported their confidence levels; they made JOL judgment after the recall phase and they asked to indicate FOK judgment before the recognition phase. Additionally, the participant group that was provided with cues presented with the same sentences that are presented in the study phase with a blank in place of a target word, and the participant group that provided feedback given fallacious feedback about the last part of the study indicating that they provided different answers to some of the questions than the majority of the participants that participated to the task. The results showed that the words in the NRP category were remembered more than the RP- words both in the recall and recognition stages, revealing that RIF manipulation worked. Meanwhile, the JOL judgments for question categories did not significantly differ. However, significantly higher FOK judgments for NRP items were observed than RP- items. Furthermore, no significant effect of testing conditions on recall and recognition performance was observed. However, a significant effect of testing conditions on participants' JOL judgments was observed. On the other hand, results indicate that testing condition has no significant effect on FOK judgments. Moreover, results indicated that participants' retrospective memory judgment scores significantly increased after the retrieval practice. Also, results indicated that testing condition has no significant effect on retrospective memory judgments and prospective memory judgments. The results were thoroughly examined and discussed concerning in relation the pertinent literature.

4.2. Discussion of the Results for Recall Performance

Concerning participant's recall performances, results indicated significant difference between item types (NRP, RP-, and RP+). Thus, while RP+ items are more correctly recalled than both NRP and RP- items, NRP items are more correctly recalled than RP- items. Therefore, these results were compatible with the fundamental findings of the RIF effect, indicating higher recall rates for NRP items than RP- items (Anderson et al., 1994; Saunders & MacLeod, 2002). However, our results also indicated that incorrect recall rates of NRP items were significantly higher than the RP- items. Generally, previous studies did not report the recall rates for incorrect items. Considering that the number of items in the NRP category (20 items) in our study was more than the number of items in the RP+ (10 items) and RP- (10 items) categories, it can be thought that the difference in incorrect recall rates is due to difference in the number of items in the categories.

When results regarding the effect of testing conditions on recall performances are examined, results indicate that the presence of cues did not affect recall performance. This finding contradicts the general idea of Tulving and Osler (1968), who indicate that stored information is facilitated by retrieval cues. Furthermore, Perfect et al., (2002) compared explicit memory tasks (cued-recall) with implicit tasks. They concluded the RIF effect is present in the cued recall condition when it fails to emerge for implicit tasks. Thus, by finding a RIF effect in cued and free recall conditions (both explicit memory tasks) our findings aligned with Perfect et al. (2002). On the other hand, this finding contradicts the previous finding by Perfect et al. (2004) that the RIF effect was robust when the cues used in the practice and test phases were the same. However, in this study, they compared category cue (provided with words) condition with face cues (with a picture of a person's face) and face plus category cue joint conditions, concluding robust RIF effect for category cue compared to face cue and face plus category joint cue. Although the cues in our study were the same as in the practice and test phases, it is thought that the difference resulted from the type of cues presented.

Furthermore, certain authors have contended that the evaluation of the RIF effect was determined by quantifying the inhibition displayed by individual participants, achieved by subtracting the performance on RP- items from NRP items (see Carroll et al., 2007; Perfect et al., 2002; Saunders, 2002). Although standard deviation exceeding means in our results may indicate floor effect for this variable, given that the inhibitory model of RIF posits that any forgetting should manifest uniformly across all testing conditions, our results corroborate this model by demonstrating a lack of significant disparity in the RIF effect across various test conditions. Additionally, the foundation of the inhibitory account of RIF suggests that retrieval cues during the practice phase create competition that produces inhibition. Consequently, this inhibition might lead to forgetting independently from the specific cues used. Thus, the RIF effect might occur regardless of the presence of the cues, pointing to the cue-independent nature of the RIF effect by supporting the inhibition account of RIF. Therefore, we concluded that the lack of significant disparity between the degree of inhibition of participants in the cued and non-cued conditions, supports the idea of the cue-independent nature of RIF, providing support for the inhibitory account.

4.3. Discussion of the Results for Recognition Performance

When the results of recognition performance after the misinformation phase are examined, it is found that even after the misinformation phase, NRP items are still significantly more correctly recalled than RP- items. Also, results for incorrect answers supported this finding by revealing significantly higher incorrect recognition rates for RP- items than NRP items. MacLeod and Saunders (2005) concluded that the RIF and misinformation effects rely on the same inhibitory process by observing higher recall of the NRP items than RP- items even after misinformation. Thus, our findings aligned with MacLeod and Saunders's (2005) study, indicating the continuity of the RIF in effect in the recognition phase even after the presentation of misinformation.

Moreover, our results indicated no significant effect of testing conditions on recognition performance that comes after the misinformation phase. Previous research tested the effect of misinformation on eyewitness memory in the context of cues (Frost & Weaver, 1997). They concluded that although cues increase the overall performance, the effect of misinformation is not reduced. However, they could eliminate the misinformation effect when numerous cues were presented after a delay and at a slower rate. Thus, it is indicated that when appropriate conditions are met in the encoding and retrieval phases, the misinformation effect can be eliminated. However, presenting numerous event cues was ineffective regarding the misinformation effect (Frost & Weaver, 1997). In our study, event cues are presented solely, and then the misinformation phase is presented before the recognition phase. Since our results did not indicate any significant difference between cued and non-cued testing conditions in terms of recognition performance, it is thought that the relationship between event cues and the misinformation effect may have been the reason behind it. It is known that, generally, the recall process is facilitated by retrieval cues (Tulving & Osler, 1968). In this context, if the presented cues reduce the effect of misinformation, it would be possible for participants to recall the original information (without the misinformed part) better, and better recall/recognition rates would have been observed. In other words, if the cues in our study were effective on the misinformation effect, it would be possible to observe the intergroup differences between the correct recognition rates of participants in the cued and non-cued conditions. Consequently, the ineffectiveness of cues on the misinformation effect and that the misinformation phase was presented before the recognition phase may have prevented the performance-enhancing effect of providing cues from occurring in the recognition phase. To sum up, the lack of influence of cues on the misinformation effect may have prevented the emergence of intergroup differences in the recognition phase.

4.4 Discussion of the Results for Metacognitive Judgments

A significantly positive correlation is observed between first and second retrospective memory judgment, prospective memory judgment, JOL, and FOK judgments. These results pointed out that there is a robust internal consistency in

participants' metacognitive evaluations. The literature also emphasizes the controversies around the issue of whether or not metacognitive judgments are processes that are domain-specific or domain-general. Domain-general processes pertain to common mechanisms employed across various tasks, whereas domain-specific processes involve distinct mechanisms that differ based on the task at hand (Irak et al., 2024; Mazancieux et al., 2020). The target accessibility theory of FOK and the retrieval fluency hypothesis of JOL are domain-general approaches that are fundamentally similar. These theories suggest that accessing more information more quickly will result in higher confidence about that item being correctly recalled or recognized later (Dunlosky & Metcalfe, 2009). Henceforth, by revealing a significantly positive correlation between metacognitive judgments throughout the RIF effect, our study confirms the findings regarding domain-general processes, which indicate that metacognitive judgments are operated by common processes, regardless of the type of task. If regionally specific processes underline the reflections in metacognitive judgments, as suggested by domain-specific approaches, it would not be possible to see strong correlations between metacognitive judgments obtained at different stages of the experimental task.

However, previous findings indicate that no correlation was found between the JOL and FOK judgments (Irak et al., 2023; Mazancieux et al., 2020; Siedlecka et al., 2016; Souchay et al., 2004; Souchay & Isingrini, 2012). The main difference in Souchay's studies is that those studies focus on measurements made with elderly individuals. Although they report that age is controlled in their correlation analysis, they also analyze metacognitive judgments' accuracy (gamma correlation). On the other hand, other studies share some differences with our study regarding the type of experimental task used and testing conditions. For instance, the experimental task used in our study was an episodic memory task that required participants to remember the words alongside the plot details rather than just words. However, Mazancieux et al. (2020) examined the correlation between episodic FOK judgments and semantic FOK judgments, episodic retrospective confidence judgments, and semantic retrospective confidence judgments in memory tasks that contain semantic memory questions and word pairs. Although Irak et al. (2023) examined the correlation between episodic and semantic FOK and retrospective confidence judgments with each other, they examined metacognitive judgments in an ERP study using an experimental task similar to

Mazancieux et al. (2020). Additionally, Siedlecka et al. (2016) obtained the metacognitive judgments of the participants (in a task in which they used various anagrams) at different stages of the study and compared them between groups. Thus, those studies highlighted the domain-specific aspects of metacognitive judgments by revealing non-significant correlations between some aspects of the metacognitive judgments.

It was stated that the human brain could include both domain-specific and domain-general neural signals, and how they interact may depend on the kind of memory task performed (Irak et al., 2023; Mazancieux et al., 2020). Relevantly, our findings support the existence and continuity of the RIF effect that activates the inhibitory mechanism at all stages of the experimental task. Thus, this finding may imply the high activation of the same cognitive mechanisms throughout the experimental task. Additionally, since both JOL and FOK judgments rely on similar monitoring and control processes, it is thought that regardless of the participants' accuracy of their assessments, participants who are judging their likelihood of correctly remembering information during recall with higher scores might also have a similar understanding of their memory processes and this may lead them to provide higher FOK judgments. Implying a possibility of a positive correlation between the metacognitive judgments when convenient task conditions that require the operation of the same cognitive mechanisms are met. Therefore, it can be thought that the continuation of the consistency between metacognitive judgments in our study may have been ensured by activating the same cognitive mechanisms that support the domain-general account of metacognitive judgments.

Our results regarding the JOL judgments of item types indicate no significant difference, although its recall performance reveals a significant difference. It is known that retrieval practice proved to be highly beneficial in enhancing learning (Karpicke & Blunt, 2011). Thus, it can be thought that individuals may have failed to separate the practiced materials, and inconsistency between memory performance and JOL judgments for practiced items could be a sign of failing metacognitive monitoring. However, results also indicate significantly higher JOL judgments for correctly answered questions than incorrectly answered ones. The fact that participants expressed higher JOL judgments for the questions they answered correctly may mean that they could distinguish correct and incorrect answers and make an

evaluation accordingly. Hence, being able to distinguish between correct and incorrect answers successfully may be an indicator that metacognitive monitoring processes are successful. Furthermore, results indicate a significant effect of testing groups on participants' JOL judgments. However, the mean differences indicated lower JOL judgments for participants when cues were presented. Providing the sentences presented during the study phase as cues reduced the participants' JOLs. The cue utilization approach of JOL indicates that JOL judgments are made based on information availability and various cues. Using relevant cues to form JOL judgments is a common process that operates independently from the type of task. Moreover, the types of cues that individuals use are distinguished as intrinsic (related to the material itself), extrinsic (related to the study conditions), and mnemonic (related to the subjective experience) cues (Koriat, 1997). Our finding supports this approach by showing that JOLs are sensitive to cues. More specifically, receiving external cues may lead individuals to interpret the material as difficult and think they need help remembering it. Therefore, their confidence in the learned item may be influenced by their subjective interpretation of their learning experience, leading to a decrease in individual JOLs.

While most memory theories propose that cues from a previously studied list would generally improve recall, it can also act as a self-limiting process, hindering access to specific categories during testing if competing category cues are employed to recall other information (Rhodes & Castel, 2008). Therefore, the results mentioned in our study are further discussed within this framework. Rhodes and Castel (2008) tested participants' memory predictions in part-set cue and free recall conditions. They concluded that although participants' overall predictions of memory performance were similar, it is highlighted that exposure to the part-set cues that resulted in diminished accessibility impairs recall performance compared to free recall. However, it is also highlighted that since no difference was observed in participants' overall predictions, they failed to accurately detect the interference caused by part-set cues.

On the other hand, in Maki's (1999) study, participants initially studied pairs of three-digit numbers (the cue) and nouns (the target). They then repeated this process and again studied digit-noun pairs, with some pairs being repeated with different targets and others being entirely new. Later, stimuli from the initial list were presented again, and participants were asked to make JOL judgments for each target. The

findings revealed that JOLs were influenced by interference, and participants could detect this interference, as lower predictions were observed for cues studied with a different target in the second phase compared to pairs entirely new across both study phases. Additionally, Tauber and Rhodes (2010) concluded that when JOLs for long and short word lists are compared, lower JOLs for longer are observed. They concluded that participants' predictions of memory performance are influenced by the number of items to be remembered. When the item quantity of our study is considered, compared to word pairs or word lists, the presentation of sentences containing varying amounts of words within the plot can indicate large quantities of to-be-remembered materials. Therefore, it can be speculated that in our study, presenting study materials that are likely to cause interference as cues and presenting to-be-remembered items in large quantities may have resulted in participants in the condition where the cues are provided having lower JOLs compared to participants in the free recall condition.

Furthermore, a comparison of gamma correlations according to the testing groups revealed moderate to high relative accuracy of JOL among four testing conditions, and the difference between the mean Gamma correlations for JOL across testing conditions was noticeable. However, statistical analysis revealed no significant differences between them. In the beginning, there were 26 participants in the cue/feedback present condition, 11 in the cue/feedback absent condition, 34 in the cue present feedback absent condition, and 7 in the cue absent/ feedback present condition. Therefore, the unequal number of people in the four test groups was thought to have prevented the possible significance of mean differences from emerging. Thus, the statistical analysis was carried out again, firstly by eliminating cue absent feedback present condition and then by eliminating cue/feedback absent condition to control the possible effects of an imbalanced number of group participants. However, results remained non-significant either way. Therefore, it is thought that there may be other reasons for significant differences not to be observed. For instance, since larger sample sizes increase the ability to detect true effects, the overall sample size in our study may still be insufficient to detect significant differences. Also, the effect size of the differences between conditions may be too small, and there could be high variability within groups. Hence, even if the means appear different, the variability within each condition might be large enough to mask these differences.

When the results related to the FOK judgments are examined, results indicate a significant difference between the FOK judgments of item types, highlighting the lower FOK judgments for non-practiced items from the practiced category. The results were in line with the memory performance. Although no prior research has come to the attention of the relationship between RIF and FOK judgments, our study sheds some light on the subject by revealing the significant difference in the FOK judgments for item types. The target accessibility theory of FOK suggests that quickly accessing information will increase metacognitive judgments (Dunlosky & Metcalfe, 2009). Consequently, access to rapid information provided through a retrieval practice phase in the RIF paradigm could affect the FOK judgments. Thus, by highlighting the significantly lower FOK judgments for non-practiced items from the practiced category (items that are thought to be inhibited), our findings support FOK's target accessibility hypothesis. Also, Carroll et al. (2017) reported that high JOL judgments were expressed for practiced items, suggesting that the process was monitored by metacognition. Since both JOL and FOK rely on the same monitoring and control processes by highlighting the significantly lower FOK judgments and recognition rates for not practiced items from the practiced category, our results were in line with the findings of Carroll et al. (2017).

On the other hand, although the RIF process is thought to be an unconscious process, providing the lowest FOK judgments for the word group with the lowest recognition rate indicates that the participants were successful in distinguishing item types, indicating successful metacognitive monitoring and suggesting that the process may not be completely unconscious. Furthermore, it is observed that higher FOK judgments are expressed for correctly answered questions compared to the incorrectly answered questions. This result implies that participants noticed their correct and incorrect answers and organized their evaluations accordingly. Hence, distinguishing between correct and incorrect answers may indicate a successful metacognitive monitoring process.

Moreover, results indicated no significant effect of testing conditions on FOK judgments. Normally, related to the cue familiarity hypothesis of FOK, it could be expected that retrieval cues may increase familiarity, resulting in enhancement of FOK judgments. However, the misinformation phase of the study is presented after the cues are provided. Thus, there is a possibility that accurate memories supported by the cues

may be diminished or overridden by the misinformation that creates memory distortion. On the other hand, Enam and McDonough (2023) pointed out the effect of cues on FOK judgments. However, in their research, font size (large vs small) was used as a perceptual cue, and level of processing (deep vs shallow) was used as a conceptual cue. They concluded that large font size only increased FOK judgments when processed at a shallow level. In our study, where event cues were used, the effect of cues was not found. Thus, when Enam and McDonough's study is considered, it can be thought that the effect of cues on FOK judgments also differs depending on the type of cue. Furthermore, the cues usually used to enhance the recall of accurate information could improve memorability and increase the sense of knowing by providing additional target information. However, in our study, no effect of cues on recall is observed (these findings were discussed above). Thus, the lack of effect of the cues on the recall may have led to memorability not being affected, and the effect of provided cues may not have been observed on FOK judgments.

Also, results indicated that second retrospective memory judgments taken after the retrieval practice are significantly higher than the first retrospective memory judgment. Retrieval practice may increase the participants' retrospective memory judgments. Thus, this finding was compatible with the literature indicating that high exposure to information and repetition can increase confidence judgments (Dunlosky & Metcalfe, 2009). On the other hand, results also indicated no effect of feedback on prospective memory judgment. However, previous studies highlighted those individuals receiving negative feedback regarding memory performance reported decreased confidence. Thus, feedback affects memory's self-credibility, altering overall memory confidence (Greenspan & Loftus, 2020). However, these studies indicate a relationship between feedback and memory confidence, but they have some common points that our study does not share. For instance, the feedback was more convincing; researchers provided it to participants. Thus, knowing that the source of the feedback is the researchers may have increased credibility by giving the impression that feedback comes from an expert and illuminating the relationship between feedback and memory confidence. More importantly, previous studies continued to collect data for a while after providing feedback rather than providing feedback at the end of the study. This issue may have affected performance by causing participants to evaluate all their judgments in line with the

feedback they received. Therefore, in our study, providing feedback via computer may have had low credibility as it did not provide concrete information about the source of the feedback. Also, the fact that feedback was provided at the end of the study may not have given enough time for participants to assess their previous performance. For these reasons, it can be thought that the relationship between feedback and memory confidence could not be stressed in our study.

Lastly, comparing gamma correlations of FOK (accuracy) according to the testing groups revealed low to moderate relative accuracy. Higher relative accuracy is observed in the cue present/feedback absent condition compared to the cue/feedback present and cue/feedback absent condition. However, despite observing notable differences in mean Gamma correlations across the four testing conditions, post hoc comparisons revealed no significant differences other than the mentioned testing conditions. As mentioned in the JOL accuracy results, sample size and statistical power, effect size, and within-group variability may have obstructed the observation of further group differences in FOK accuracy.

On the other hand, FOK's relative accuracy was lower than the JOL's accuracy. Previous studies indicate that higher JOL accuracy is achieved through practice (Mazzonni & Nelson, 1995). Also, Dunlosky and Metcalfe (2009) highlighted that suggestive questions can make people think that misleading information is true, which is a sign of failing metacognitive monitoring. Thus, higher relative accuracies for JOL judgments than FOK judgments may have shown us different efficacy of both practice and misinformation on metacognitive judgments. More clearly, the retrieval practice phase presented before the JOL judgments may have increased the accuracy. At the same time, the misinformative questions presented before the FOK judgments were taken had the opposite effect. Future studies may shed light on the subject by examining the retrieval practice and misinformative questions with metacognitive judgments in separate studies. For instance, to test the sole effect of retrieval practice in an experimental study, while the participants in the experimental group participated in the retrieval practice, the participants in the control group could relate to the filler task rather than seeing retrieval practice materials, and both groups subsequently could report their JOL judgments about the materials.

In conclusion, the results shed some light on this paradigm's metacognitive monitoring and control processes. The results highlighted that the RIF process, which

is thought to be unconscious, may not be completely unconscious, and actions in the RIF paradigm are successfully monitored by metacognition.

4.5 Strengths and Limitations

Several limitations and strengths can be cited in the present study. Firstly, conducting the study online via computer resulted in a lack of control over the study environment and a lack of real-time monitoring. For this reason, it was not checked whether the participants understood the instructions, and if they had questions, they were not answered during the study. Secondly, the study took between 45 and 50 minutes to complete, depending on the progress of the participants, which may have resulted in possible fatigue. This situation may affect the participants' attention, motivation, and performance, particularly especially toward the end of the study. On the other hand, there are also several strengths of this study. The first one is to best our knowledge, this is the first study to examine metacognitive judgments under the RIF paradigm and misinformation effect. Secondly, conducting this study online also provided access to a more diverse sample from different locations and a larger sample size. Lastly, stabilizing the time that the study materials remained on the screen (except for the recall phase) ensured continuity in the experimental procedure across participants by reducing the variability in participant performance due to timing.

4.6. Contributions to the Field and Recommendations for Future Research

This study contributes to the field by providing a comprehensive perspective on metacognitive monitoring and control processes by examining metacognitive judgments regarding memory errors in the context of external variables (e.g., cues and feedback). On the other hand, although it was observed that giving clues did not affect metacognitive judgments, it caused a decrease in JOL judgments. Therefore, future studies can better compare different cues to understand the sensitivity of metacognitive judgments about them. In addition, in future studies, different types of feedback (such as visual and auditory) can be included in different ways, and whether there are any changes in metacognitive judgments can be examined through behavioral and brain

imaging studies. Thanks to these studies, more detailed information can be obtained about the basis of diseases in which memory errors and long-term memory problems occur (such as Alzheimer's, Schizophrenia, post-traumatic stress disorder, and types of addiction). Finally, this information may contribute to developing cognitive rehabilitation programs for the mentioned diseases.



REFERENCES

- Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(5), 1063–1087. <https://doi.org/10.1037/0278-7393.20.5.1063>
- Arbuckle, T. Y., & Cuddy, L. L. (1969). Discrimination of item strength at time of presentation. *Journal of Experimental Psychology*, 81(1), 126–131. <https://10.1037/h0027455>
- Ashcraft, M. H., & Radvansky, G. A. (2018). *Cognition*. Pearson Education.
- Butler, K. M., Williams, C.C., & Maki, R. H. (2001) A Limit on Retrieval-induced Forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 27(3), 1214-9. <https://10.1037//0278-7393.27.5.1314>.
- Butler, A. C., Karpicke, J. D., & Roediger, H. L. (2008). Correcting a metacognitive error: Feedback increases retention of low-confidence correct responses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(4), 918–928. <https://10.1037/0278-7393.34.4.918>
- Carroll, M., Campbell-Ratcliffe, J., Murnane, H., & Perfect, T. (2007). Retrieval-induced forgetting in educational contexts: Monitoring, expertise, text integration, and test format. *European Journal of Cognitive Psychology*, 19(4-5), 580–606. <https://10.1080/09541440701326071>
- Challies, D. M., Hunt, M., Garry, M., & Harper, D. N. (2011). Whatever Gave You That Idea? False Memories Following Equivalence Training: A Behavioral Account of the Misinformation Effect. *Journal of the Experimental Analysis of Behavior*, 96(3), 343–362. <https://10.1901/jeab.2011.96-343>
- Connor, L. T., Dunlosky, J., & Hertzog, C. (1997). Age-related differences in absolute but not relative metamemory accuracy. *Psychology and Aging*, 12(1), 50–71. <https://10.1037/0882-7974.12.1.50>

- Conway, M. A., Singer, J. A., & Tagini, A. (2004). The Self and Autobiographical Memory: Correspondence and Coherence. *Social Cognition*, 22(5), 491–529. <https://10.1521/soco.22.5.491.50768> 10.1521/soco.22.5.491.50768
- Conway, M. A., Justice, L. V., & D'Argembeau, A. (2019). The Self-Memory System Revisited: Past, Present, and Future. In J. H. Mace (Ed.), *The Organization and Structure of Autobiographical Memory* (pp. 76-134). Oxford University Press.
- Dunlosky, J., & Metcalfe, J. (2009) *Metacognition: A Textbook for Cognitive, Educational, Life Span & Applied Psychology*. (Evans, E., Ed.) Sage Publications
- Enam, T & McDonough, I. M. (2024). Integrating multiple cues in metamemory: using the illusory effect of font size and level of processing to inform FOK judgments. *Metacognition and Learning*, 19,169–188. <https://doi.org/10.1007/s11409-023-09367-6>
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906–911. <https://doi.org/10.1037/0003-066X.34.10.906>
- Frost, P., & Weaver, C. A. (1997). Overcoming Misinformation Effects in Eyewitness Memory: Effects of Encoding Time and Event Cues. *Memory*, 5(6), 725–740. <https://10.1080/741941545>
- Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, 24(1), 13–23. [https://10.1016/s0160-2896\(97\)90011-8](https://10.1016/s0160-2896(97)90011-8)
- Greenspan, R. L., & Loftus, E. F. (2020). Eyewitness confidence malleability: Misinformation as post-identification feedback. *Law and Human Behavior*, 44(3), 194–208. <https://doi.org/10.1037/lhb0000369>
- Hart, J. T. (1966). Methodological note on feeling-of-knowing experiments. *Journal of Educational Psychology*, 57(6), 347–349. <https://10.1037/h0023915>
- Hart, J. T. (1967). Memory and the memory-monitoring process. *Journal of Verbal Learning and Verbal Behavior*, 6(5), 685–691. [https://10.1016/s0022-5371\(67\)80072-0](https://10.1016/s0022-5371(67)80072-0)
- Hisli, N. (1988) Beck Depresyon Envanterinin Gecerligi Uzerine bir Calisma. *Journal of Psychology*, 6, 118-122.

- Hollins, T. J., & Weber, N. (2016). Monitoring and Regulation of Accuracy in Eyewitness Memory: Time to Get Some Control. In J. Dunlosky & S. U. K. Tauber (Eds.), *The Oxford Handbook of Metamemory* (pp. 171-195). Oxford University Press.
- Irak, M. (2005) Üst- Biliş Mi? Yönetici İşlevler Mi?: Bilme Hissinin Nöropsikolojik Testlerle Ölçülen Dikkat Süreçlerinden Yordanması. *Türk Psikoloji Dergisi*. 20(56), 97-116.
- Irak, M., Soylu, C., Turan, G., & Çapan, D. (2019). Neurobiological basis of feeling of knowing in episodic memory. *Cognitive Neurodynamics*. <https://10.1007/s11571-019-09520-5>
- Irak, M., Soylu, C., & Yavuz, M. (2023). Comparing event-related potentials of retrospective and prospective metacognitive judgments during episodic and semantic memory. *Scientific Reports*, 13(1), 1949.
- Irak, M., Topçuoğlu, V., Duman, T. N., Akyurt, S., Yılmaz, İ., & Pala, İ. Y. (2024). Investigating Retrospective and Prospective Metamemory Judgments During Episodic Memory in Patients With Obsessive-Compulsive Disorders. *Behavior Therapy*, 55(2), 277-291.
- Karpicke, J. D., & Blunt, J. R. (2011). Retrieval Practice Produces More Learning than Elaborative Studying with Concept Mapping. *Science*, 331(6018), 772–775. <https://10.1126/science.1199327>
- Khera, T. & Rangasamy, V. (2021). Cognition and Pain: A Review. *Frontiers Psychology*. 12:673962. <https://doi.org/10.3389/fpsyg.2021.673962>
- Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review*, 100(4), 609–639.
- 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. <https://doi.org/10.1037/0096-3445.126.4.349>
- Koriat, A., Sheffer, L., & Ma'ayan, H. (2002). Comparing objective and subjective learning curves: Judgments of learning exhibit increased underconfidence with practice. *Journal of Experimental Psychology: General*, 131(2), 147–162. <https://10.1037/0096-3445.131.2.147>
- Kornell, N., & Rhodes, M. G. (2013). Feedback reduces the metacognitive benefit of tests. *Journal of Experimental Psychology: Applied*, 19, 1– 13.

- Loftus, E. F., Donders, K., Hoffman, H. G., & Schooler, J. W. (1989). Creating new memories that are quickly accessed and confidently held. *Memory & Cognition*, 17(5), 607–616. [https://10.3758/bf03197083](https://doi.org/10.3758/bf03197083)
- Loftus, E.F. & Pickrell, J.E. (1995) The formation of false memories. *Psychiatric Annals*, 25, 720-725.
- MacLeod, M. D., & Saunders, J. (2005). The Role of Inhibitory Control in the Production of Misinformation Effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(5), 964–979. [https://10.1037/0278-7393.31.5.964](https://doi.org/10.1037/0278-7393.31.5.964)
- Maki, R. H. (1999). The roles of competition, target accessibility, and cue familiarity in metamemory for word pairs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(4), 1011–1023. [https://10.1037/0278-7393.25.4.1011](https://doi.org/10.1037/0278-7393.25.4.1011)
- Mazancieux, A., Dinze, C., Souchay, C., & Moulin, C. J. A. (2020). Metacognitive domain specificity in feeling-of-knowing but not retrospective confidence. *Neuroscience of Consciousness*, (1). [https://10.1093/nc/niaa001](https://doi.org/10.1093/nc/niaa001)
- 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. [https://10.1037/0278-7393.21.5.1263](https://doi.org/10.1037/0278-7393.21.5.1263)
- Melton, A. W., (1963). Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning and Verbal Behavior*, 2, 1–21.
- Metcalfe, J., & Shimamura, A. P. (Eds.). (1994). Metacognition: Knowing about knowing. The MIT Press. <https://doi.org/10.7551/mitpress/4561.001.0001>
- Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, 84, 93–116.
- Nelson, T.O. & Narens, L. (1990). Metamemory: A theoretical framework and new finding.” Bower, G. (Ed.), *The Psychology of Learning and Motivation* (26), New York: Academic Press.
- Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1–25). The MIT Press.

- Nelson, T. O. (2001). Metamemory. *Psychology of International Encyclopedia of the Social & Behavioral Sciences*, 9733–9738. <https://10.1016/b0-08-043076-7/01516-3>
- Perfect, T. J., Moulin, C. J. A., Conway, M. A., & Perry, E. (2002). Assessing the inhibitory account of retrieval-induced forgetting with implicit-memory tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(6), 1111–1119. <https://10.1037/0278-7393.28.6.1111>
- Perfect, T. J., Stark, L.-J., Tree, J. J., Moulin, C. J. A., Ahmed, L., & Hutter, R. (2004). Transfer appropriate forgetting: The cue-dependent nature of retrieval-induced forgetting. *Journal of Memory and Language*, 51(3), 399–417. <https://10.1016/j.jml.2004.06.003>
- Pintrich, P. R. (2002). The Role of Metacognitive Knowledge in Learning, Teaching, and Assessing. *Theory Into Practice*, 41(4), 219–225. https://10.1207/s15430421tip4104_3
- Ren, Z., Liu, W., & Wang, L. (2023) Can Feedback Boost Prospective Memory Performance? *American Journal of Psychology*. 136 (1), 71-88.
- Rhodes, M. G., & Castel, A. D. (2008). Metacognition and part-set cuing: Can interference be predicted at retrieval? *Memory & Cognition*, 36(8), 1429–1438. <https://10.0.14.174/mc.36.8.1429>
- Rhodes, M. G. (2016). Judgments of Learning: Methods, Data, and Theory. In J. Dunlosky & S. U. K. Tauber (Eds.), *The Oxford Handbook of Metamemory* (pp. 65-80). Oxford University Press.
- Salovich, N. A., & Rapp, D. N. (2021). Misinformed and unaware? Metacognition and the influence of inaccurate information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 47(4), 608–624.
- Saunders, J., & MacLeod, M. D. (2002). New evidence on the suggestibility of memory: The role of retrieval-induced forgetting in misinformation effects. *Journal of Experimental Psychology: Applied*, 8(2), 127–142. <https://10.1037/1076-898x.8.2.127>
- Schacter, D. L. (1999) The seven sins of memory. *Insights from psychology and cognitive neuroscience*. 54(3), 182-203. <https://10.1037//0003-066x.54.3.182>
- Schwartz, B. L., & Metcalfe, J. (1992). Cue familiarity but not target retrievability enhances feeling-of-knowing judgments. *Journal of Experimental Psychology:*

- Learning, Memory, and Cognition, 18(5), 1074–1083.
<https://doi.org/10.1037/0278-7393.18.5.1074>
- Shaw, J. S., Bjork, R. A., & Handal, A. (1995) Retrieval-induced forgetting in an eyewitness memory paradigm. *Psychonomic Bulletin & Review*, 2, 249-253
- Siedlecka, M., Paulewicz, B., & Wierzchoń, M. (2016). But I Was So Sure! Metacognitive Judgments Are Less Accurate Given Prospectively than Retrospectively. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.00218>
- Sitzman, D. M., Rhodes, M. G., & Tauber, S. K. (2013). Prior knowledge is more predictive of error correction than subjective confidence. *Memory & Cognition*, 42(1), 84–96. <https://doi.org/10.3758/s13421-013-0344-3>
- Sitzman, D. M., Rhodes, M. G., & Kornell, N. (2016). The influence of feedback on predictions of future memory performance. *Memory & Cognition*, 44(7), 1102–1113. <https://doi.org/10.3758/s13421-016-0623-x>
- Son, L. K., & Metcalfe, J. (2005). Judgments of learning: Evidence for a two-stage process. *Memory & Cognition*, 33(6), 1116–1129. <https://doi.org/10.3758/bf03193217>
- Souchay, C., Isingrini, M., Clarys, D., Taconnat, L., & Eustache, F. (2004). Executive Functioning and Judgment-of-Learning versus Feeling-of-Knowing in Older Adults. *Experimental Aging Research*, 30(1), 47–62. <https://doi.org/10.04.56/03610730490251478>
- Souchay, C., & Isingrini, M. (2012). Are feeling-of-knowing and judgment-of-learning different? Evidence from older adults. *Acta Psychologica*, 139(3), 458–464. <https://doi.org/10.1016/j.actpsy.2012.01.007>
- Squire, L. (1986). *Mechanisms of memory*. *Science*, 232(4758), 1612–1619. <https://doi.org/10.1126/science.3086978>
- Squire, L. R. (1992). Declarative and Nondeclarative Memory: Multiple Brain Systems Supporting Learning and Memory. *Journal of Cognitive Neuroscience*, 4(3), 232–243. <https://doi.org/10.1162/jocn.1992.4.3.232>
- Storm, B. C., & Levy, B. J. (2012). A progress report on the inhibitory account of retrieval-induced forgetting. *Memory & Cognition*, 40(6), 827–843. <https://doi.org/10.3758/s13421-012-0211-7>

- Sutton, J., Harris, C. B., & Barnier, A. J. (2010). 14 Memory and Cognition. In S. Radstone & B. Schwarcz (Eds.), *Memory: Histories, Theories, Debates* (pp. 209–226). Fordham University Press.
- Tauber, S. K., & Rhodes, M. G. (2010). Does the amount of material to be remembered influence judgements of learning (JOLs)? *Memory*, 18(3), 351–362. <https://10.1080/09658211003662755>
- Tulving, E., & Osler, S. (1968). Effectiveness of retrieval cues in memory for words. *Journal of Experimental Psychology*, 77(4), 593–601. <https://10.1037/h0026069>
- Tulving, E. (1984). How Many Memory Systems Are There? *American Psychological Association*. 40(8). 385-398.
- Ulusoy M, Şahin N, Erkman H. (1998) Turkish version of the beck anxiety inventory: psychometric properties. *J.Cogn Psychother: An Int Quarterly* 12, 28-35.
- Vortac, O. U., Edwards, M. B., & Manning, C. A. (1995). Functions of External Cues in Prospective Memory. *Memory*, 3(2), 201–219. <https://10.1080/09658219508258966>
- Wells, G. L., & Bradfield, A. L. (1998) Good, you identified the suspect: Feedback to eyewitnesses distorts their reports of the witnessing experience. *Journal of Applied Psychology*, 83, 360-376
- Wells, G. L., & Loftus, E. F. (2013). Eyewitness memory for people and events. In R. K. Otto & I. B. Weiner (Eds.), *Handbook of psychology: Forensic psychology* (pp. 617–629). John Wiley & Sons, Inc.
- Zhang, F., Zhang, X., Luo, M., & Geng, H. (2016). The Effects of Feedback on Memory Strategies of Younger and Older Adults. *PLOS ONE*, 11(12). <https://10.1371/journal.pone.01688>
- Zhou, L., Lu, J., & Dong, L. (2016). Age-Related Differences in Metamemory Accuracy among the Elderly: The Effect of Declining Inhibitory Function. *Journal of Adult Development*, 24(1), 48–57. <https://10.1007/s10804-016-9246-5>