THE EFFECT OF TIMING AND STIMULUS PROPERTIES

ON METACOGNITIVE JUDGMENTS

DİLAY ZEYNEP KARADÖLLER ASTARLIOĞLU

BOĞAZİÇİ UNIVERSITY

2016

THE EFFECT OF TIMING AND STIMULUS PROPERTIES ON METACOGNITIVE JUDGMENTS

Thesis submitted to the

Institute for Graduate Studies for Social Sciences

in partial fulfillment of the requirements for the degree of

Master of Arts

in

Psychological Sciences

by

Dilay Zeynep Karadöller Astarlıoğlu

Boğaziçi University

2016

DECLARATION OF ORIGINALITY

I, Dilay Zeynep Karadöller Astarlıoğlu, certify that

- I am the sole author of this thesis and that I have fully acknowledged and documented in my thesis all sources of ideas and words, including digital resources, which have been produced or published by another person or institution;
- this thesis contains no material that has been submitted or accepted for a degree or diploma in any other educational institution;
- this is a true copy of the thesis approved by my advisor and thesis committee at Boğaziçi University, including final revisions required by them.

Signature..

Date ......2.2.2016......

ABSTRACT

The Effect of Timing and Stimulus Properties on Metacognitive Judgments

The present thesis tested the effect of partial information along with the effect of timing on feeling of knowing (FOK) judgments to provide a comprehensive account of variables influencing FOK strength and accuracy. In the present study, partial information was conceptualized based on Koriat (1994)'s argument such that partial information is something that comes to mind in thinking about the target and is the 'glimpse' of the information related to the target. To manipulate levels of partial information available at retrieval, we used a cued recall task and paired medium frequency words with words that either had large or small association sets and weak or strong primary associates, respectively. Moreover, we also manipulated whether FOK judgments were given under time constraints or in a self-paced fashion. We conducted three successive experiments. Most critically we found that people's FOK judgments were affected by both timing and level of partial information manipulations. In particular, participants gave higher FOK judgments for correctly recalled trials when they had enough time to evaluate their judgments regardless of the association level of the words. However, for incorrectly recalled items higher judgments were given for the targets with strong levels of association and in the self-paced time conditions. Results of FOK accuracy did not differ across timing manipulation and word type. These findings suggest that both strong level of partial information of a given word and self-paced timing for making a FOK judgment increases FOK judgment magnitude.

ÖZET

Zamanlamanın ve Uyarıcı Özelliklerinin Üstbilişsel Yargılara Etkileri

Bu tezin amacı bir kelimenin ilgili tüm kısmi özelliklerinin ve bununla birlikte zamanlamanın, Bilme Hissi Yargısı üzerindeki etkisini aynı kapsamda incelemektir. Sonuçlar Bilme Hissi Derecesi ve bu derecenin kişinin hatırlama performansıyla uyumluluğu bağlamında incelenmiştir. Bu çalışmada kelimenin kısmi özellikleri Koriat (1994)'ın argümanı baz alınarak kurgulanmıştır. Bu argümana göre kısmi bilgi, kişinin bir kelime ile ilgili düşünürken aklına gelen bilgilerin tümünü kapsamaktadır. Bu çalışma özelinde kısmi bilgi düzeyini, güçlü yada zayıf çağrışım seti olan kelime grupları kullanarak deneysel yolla manipüle ettik. Buna ek olarak, Bilme Hissi Yargısı'nın verilme süresini, bazı katılımcılara yeterli süre, bazı katılımcılara ise kısıtlı cevaplama süresi vererek farklı iki koşulda inceledik. Art arda yaptığımız üç çalışmada bulduğumuz sonuçlara göre, insanların Bilme Hissi Yargısı hem cevaplama zamanlaması hem de kelimenin kısmi özelliklerinden etkilenmektedir. Aynı zamanda kişinin serbest hatırlama performansının da bu sonuçlar üzerinde etkisi vardır. Katılımcılar, doğru hatırladıkları kelimeler için yeterli zaman verildiğinde, kısıtlı zaman verilen duruma göre daha yüksek Bilme Hissi Yargısı'nda bulunmuşlardır. Oysa ki, hatırlayamadıkları ya da yanlış hatırladıkları kelimeler arasında, sadece yüksek kısmi bilgiye sahip kelimeler için, ve sadece yeterli cevaplama süresi verildiği koşulda yüksek Bilme Hissi Yargısı'nda bulunmuşlardır. Genel olarak sonuçlar , güçlü çağrışıma sahip kelimer için cevaplama süresi serbest bırakıldığında yüksek bilme hissi yargısında bulunulmakta olduğunu göstermektedir.

ACKNOWLEDGEMENTS

I owe a great many thanks to a great many people who supported me in making this endeavor a reality. Foremost, I would like to express my deep sense of gratitude to my thesis advisor, Assoc. Prof. Ayşecan Boduroğlu, who has always been a great source of encouragement, inspiration and motivation to me both personally and academically. Her tailored guidance, constructive comments, and friendly attitude have been very much appreciated. I feel very lucky for having a chance to work under her supervision. I would like to express my gratitude to Prof. Ali Tekcan. Since my undergraduate studies, he has always been a source of inspiration to me. I am really grateful to him in letting me work in his laboratory and believing me in accomplishing my goals. I heartily acknowledge that he has helped me to find my way in the academic life. I am very grateful to Prof. Bennett Schwartz who was always ready to help. I am really indebted to him for his bright ideas, critical feedbacks, and excitement about the work. Furthermore, I genuinely acknowledge the keen interest of Assist. Prof. Esra Mungan in this thesis. Her valuable comments and support is very much appreciated.

I am very thankful to my colleagues in assistants' office, Burcu Kaya Kızılöz, Erdem Ozan Meral, Melih Barsbey, Merve Mutafoğlu, Naziye Güneş Acar, Pınar Ersoy and Şule Pala Sağlam for their endless support and invaluable friendship. Their warm similes turned that little office into a peaceful workplace.

I am also very grateful to my dearest friends from the laboratory, Aslı Aktan Erciyes, Bilge Göz, Ezgi Mamus, Gamze Sofuoğlu, Müge Özbek, Müge Özvarol, and Semra Esen. Their friendship and genuine support helped me to endure all troubles and difficulties in my academic and personal life. I know our genuine friendship and academic collaboration will continue wherever we are in the future.

I sincerely thank to Melis Gülerdi, Selen Küçüktaş, and Zeynep Akça for their support in data collection phase of this thesis and Aysu Mutlutürk for her help in construction of the testing material. I also want to thank TUBITAK (Scientific and Technological Research Council of Turkey) for its grants during my entire academic life.

I would like to express my very profound gratitude to my parents Dilek Karadöller and Aydın Karadöller and my brother Egecan Karadöller for providing me with unfailing support and continuous encouragement throughout my years of study.

Finally, I must express my very great appreciation to my bellowed husband, Melih Astarlıoğlu, for his endless support and encouragement throughout the process of researching and writing this thesis. He has been not only a life partner but also a role model both personally and academically. This accomplishment would not have been possible that much easier without him. Thank you.

To Melih..

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

THEORETICAL BACKGROUND

One of the most interesting memory phenomena occurs when people feel that they know the answer to a question, although they cannot produce the answer. In certain situations people often feel that the answer is on the "Tip-of-the-Tongue" or may report partial information about the answer even when they cannot recall it (Brown, 1991; Schwartz & Smith, 1997). These phenomena demonstrate metamemory, the processes by which we know what we have stored in our memory. Metamemory processes are composed of two components; monitoring and control (Nelson, 1996; Nelson & Narens, 1990). The monitoring component is responsible for observing the content of memories and it allows us to make judgments about information stored in our memory, whereas, the control component is conceptualized to be responsible for the regulation of information between actual memory stores and the monitored information. Metamemory processes are most commonly studied by investigating Tip-of-the-Tongue (TOT) states, Feeling-of-Knowing (FOK), Judgments of Learning (JOL), confidence (CONF) and source judgments (Metcalfe & Dunlosky, 2009).

FOK judgments are classically defined as the subjective likelihood of future recognition for currently unrecalled items (Hart, 1965). Most early studies on FOK focused on the accuracy of FOK judgments in predicting actual memory performance (e.g. Hart, 1965; Nelson, 1984). In the subsequent years, however, researchers have focused more on determining the underlying mechanisms for FOK judgments (e.g. Koriat & Levy-Sadot, 2001; Schwartz & Metcalfe, 1992).

Nelson and his colleagues proposed that all possible explanations regarding the mechanisms of FOK judgments are grouped into two (Nelson, Gerler & Narens, 1984). These are trace access account and inferential mechanism. The trace access account assumes that people can directly monitor the target in memory even in the presence of recall failure (Koriat, 1994). On the other hand, inferential mechanisms focus on the availability of non-target information, such as semantic or episodic information related to the target (Koriat & Levy-Sadot, 2001). Research to date has provided greater support for the inferential accounts (but see Hertzog, Fulton, Sinclair, Dunlosky, 2014). Two of the most important inferential accounts of FOK are cue familiarity (Metcalfe, Schwartz & Joaquim, 1993; Reder, 1987) and target accessibility (Koriat, 1993). The cue familiarity account states that FOK is influenced by the familiarity with the memory pointer, that is, the stimulus to cue the actual memory (Metcalfe, 1993). On the other hand, target accessibility account does not focus on cues and suggests that FOK is influenced by the overall accessibility of the partial information related to the target (Koriat, 1993; 1994; 1995). Some more recent research has provided support for a more combined, hybrid account (Benjamin, 2005; Hosey, Peynircioğlu & Rabinovitz, 2009; Koriat & Levy-Sadot, 2001).

Metacognitive research basically focuses on identifying certain factors influencing metacognitive judgments, particularly magnitude of FOK judgments and accuracy of these judgments. FOK magnitude is influenced by inferential mechanisms such as cue familiarity (Koriat, 1993; Metcalfe, Schwartz, Joquaim, 1993) and target accessibility (Koriat, 1995). FOK accuracy is conceptualized as the calibration between the FOK magnitude and actual memory performance and measured by Goodman-Kruskal gamma correlations of the FOK with recognition accuracy (Gonzales & Nelson, 1996). Thus, these two different levels of judgments – FOK magnitude and FOK accuracy- are independent of each other due to scaling differences. Both FOK magnitude and FOK accuracy is influenced by accessibility of upcoming information related to the target but they differ in certain situations. For example, if the accessed information includes false memories FOK accuracy decreases as FOK magnitude increases (Koriat, 1995).

According to the accessibility model of Koriat (1993; 1994; 1995), FOK accuracy is determined not only by the accuracy of partial information related to the target, but also by the amount as well as timing of the retrieval of that information. The non-criterial recollection hypothesis extends this view further and claims that the amount of non-criterial source information influences FOK (Brewer, March, Clark-Foos & Meeks, 2010). This account defines non-criterial information as all information from the encoding context except the target itself. For instance, any information related to the target or participants' emotional state at encoding is considered as non-criterial information. In this regard, non-criterial information is rather different than partial information in that the latter is exclusively related to the target. Greatest support for the non-criterial hypothesis view comes from evidence showing that both correct and incorrect non-criterial source information contributes to FOK.

Despite the empirical support for the more classical accounts summarized above, it is also known that these accounts only explain a relatively small percentage of the variance in FOK judgments (e.g. Koriat, 1993; Koriat & Levy-Sadot, 2001; Metcalfe, Schwartz & Joaquim, 1993). Given the support that more alternative and modern views like the non-criterial hypothesis has received, it is critical to understand the properties of the target that may influence retrieval processes and consequently FOK judgments (Brewer et al., 2010; Schwartz, Pillot & Bacon, 2014; Thomas, Bulevich & Dubois, 2011; 2012).

1.1 Impact of stimulus properties on memory and FOK

Many of the episodic memory studies highlight the importance of stimulus properties in determining memory performance. For instance, when people are asked to remember certain previously studied information, people tend to hold an optimum amount of information in memory rather than the whole (Pansky & Koriat, 2004). This particular strategy may allow one to optimize the level of information held in memory. To illustrate, Pansky and Koriat (2004) presented people with either superordinate (e.g. vehicle) or subordinate (e.g. sports car) information in the study phase. They found that regardless of the initial information provided in the study phase, people's retrieval of that information converged onto the basic level (e.g. car). This finding is observed for both recall and recognition tests and also found in both immediate and delayed testing phases. The impact of the stimulus properties and more specifically of target related partial information on retrieval is illustrated in another study. Koriat and his colleagues presented Hebrew speaking participants with pseudo Somali words (i.e. one to three syllable pronounceable nonsense strings which did not relate to any Hebrew word) and their Hebrew translations. Although participants showed a decline in item recall performance after a one-week delay, their memory for the partial information identification (polarity of the words in terms of dimensions such as evaluation, potency and activity) related to these words were found to be equivalent across the week delay (Koriat, Levy-Sadot, Edry & de Marcas, 2003).

Encoding strategies have been shown to impact memory and metamemory performance. Researchers have investigated the impact of manipulations during encoding on FOK judgments (e.g. Hertzog, Dunlosky, & Sinclair, 2010; Lupker, Harbluk, & Partick, 1991; Sacher, Tacconat, & Souchay, 2009; Schacter, 1983; Tekcan & Aktürk, 2001). For instance, Boduroglu, Pehlivanoglu, Tekcan and Kapucu (2015) showed that self-referencing at encoding resulted in better memory performance and higher FOK accuracy.

Previous research also presents evidence on how partial information related to the target influence FOK judgments. Koriat (1993) demonstrated that increasing the amount of partial information retrieved by participants increased their FOK judgment magnitude (Koriat, 1993). In this study, Koriat presented participants in two different experiments either with four or five letter combination of consonants (e.g. RDFK or RDFKZ) to study, and later asked them to recall and provide a FOK judgment. He found that, regardless of the correctness of the information, increases in the amount of retrieved partial information correlated with increased FOK judgment magnitude. Similarly, Schwartz and his colleagues (2014) experimentally manipulated the amount of contextual information, regarding imaginary animals, presented to participants. In this study, contextual information included country of origin, diet and weight of the imaginary animals. During testing, in addition to contextual information, they collected the partial information about the animal’s name provided by the participants. The experimental manipulation showed that FOK strength was higher when more contextual information was provided at study. They also found a tendency for the increase in partial information along with the increases in FOK judgments, but this finding failed to reach significance.

Apart from the quantity of partial information, recent evidence suggests that the quality of partial information retrieved also influence FOK magnitude. For example, Thomas and her colleagues (2011) found that people gave significantly higher FOK judgments when they correctly remembered the valence of the target. In a subsequent study, Thomas and her colleagues (2012) found that the qualitative features of the stimuli selectively influenced FOK judgments: Questions regarding the perceptual nature of pictures or ambiguous figures did not impact FOK magnitude; however, questions regarding the semantic nature of the pictures or namable pictures increased it. All these studies stress that both quantitative and qualitative nature of the partial information and information related to the target is influential on FOK judgments.

1.2 Temporal dynamics of FOK

The time allotted to make a judgment is also known to influence FOK magnitude. According to the accessibility model (Koriat, 1993; 1994; 1995), retrieval mechanisms require time to retrieve partial information. Therefore, if a person spends more time attempting retrieval of an unrecalled item, more partial information accumulates, leading to higher FOK judgments. It is important to stress that, not only the amount of partial information but also the speed with which partial information is recovered matters for FOK judgments. In relation to this view, Thomas and her colleagues found that magnitude of FOK judgments are influenced by the time allotted for responses (Thomas, et. al., 2011). Specifically, Thomas et al. argued that restricting the timing of response to

1500 milliseconds (ms) in an episodic memory task and forcing people to respond very quickly blocked the evaluative processing of partial information, thus lowering FOK accuracy. There was no effect of the timing manipulation on FOK magnitude. Additionally, Koriat and Levy-Sadot (2001) argued that the mechanism that people rely on differs based on the timing of the FOK judgment. Specifically, during the initial stages of retrieval, people give FOK judgments based on the familiarity of the cue. However, if people were provided with sufficient time, they made their FOK judgments based on accessibility (also see Schwartz, 1993). Similarly, Benjamin (2005) manipulated cue and target duration as well as response timing in an episodic memory task. He found that under restricted time, predictions were driven by cue familiarity; however, under self-paced conditions predictions were influenced by target duration. Thus, these results reflect the usage of different mechanisms based on the timing of FOK judgment.

Overall, considering the previous findings, FOK judgments depend on two important sources. The first one is the nature of the partial information related to the target itself and the second one is the time allotted for the FOK judgment. Despite many studies focusing on either one of these issues, no study to date has provided an integrated view addressing the interaction between these two factors.

1.3 The present study

The aim of the present study was to test the effect of partial information related to target along with the effect of timing to derive a comprehensive explanation for the accuracy and magnitude of FOK judgments. In order to accomplish this aim, we manipulated two things. First of all, we manipulated the amount of partial information related to the target. We also experimentally manipulated when people could produce their FOK judgments.

In the present study, effects of partial information were conceptualized based on Koriat (1994). Koriat argued that partial information is anything that comes to mind while thinking about the target. It could be either based on solid referents such as the first letter of a word, color of the book if that is what was being retrieved, or even mere feelings related to the target. Partial information captures every single bit of information of any type and its correctness does not matter. In the present study, we experimentally manipulated the amount of partial information associated with targets by using words that had either weak or strong levels of partial information at the conceptual level.

The strength of partial information was experimentally manipulated via using words from different association sets. Free Association Norms, having high degrees of reliability (r = .89; Nelson, McEvoy & Schrieber, 2004), index the likelihood that a particular word cues another. These norms were representative of the relative accessibility of related words in one's memory. In other words, free association norms measure links that bind associated words together (Nelson, McEvoy & Schrieber, 2004). According to Nelson, McEvoy and Dennis (2000) words in free association sets could be classified into three subgroups in terms of set size (i.e. large set or small set) and strength of the primary associate (i.e. weak vs. strong). Within these subgroups, words with small set size and strong primary associates were found to have the highest interassociation compared to words with other sets (small set-weak primary and large set-weak primary). There is no word group with large set – strong primary because strong primaries require one highly dominant primary associate which in turn limits the number of other associations creating a small set. This finding presented that words with small set size and strong primary associations brought about same words across participants and words with largest associative arrays with weak primaries brought about idiosyncratic associations and did not belong to a predictable web of words. Additionally, according to Processing Implicit and Explicit Representations Hypothesis (PIER; Nelson, Schrieber & McEvoy, 1992) retrieval of target words depends on either explicit or implicit processing activities, and set size of words varies as the amount of implicit information increases. In other words, they found that with smaller set size, the likelihood of remembering the target was greater. Therefore, we conceptualized that words with small set size and strong primary associates could be regarded as having strong partial information, and words with large set size and weak primary associates as having weak partial information.

In order to manipulate the effect of timing on FOK judgments we enforced participants to make FOK judgments in a restricted time (<2000ms). Participants were either asked to respond within 2000ms or in their self-pace. Aim of this manipulation was to block possible evaluative processing that people might engage while making FOK judgments.

In order to test the effect of timing and level of partial information, the current study incorporated a 2 (levels of partial information of the stimuli: weak or strong) X 2 (time allotted for judgment: restricted or self-paced) between subjects design. This design was valid for three consecutive experiments conducted for the purpose of this thesis.

1.4 Hypotheses

Based on previous studies, we hypothesized that in a condition in which participants were given restricted time, stimuli with strong levels of partial information would create high familiarity, and thus participants would be more likely to give higher FOK judgments. However, because participants could not evaluate the upcoming partial information due to time restriction we expected them to fail in predicting their actual recognition performance by overestimating their performance compared to participants given stimuli with weak partial information. When participants study words with weak levels of partial information and respond in a restricted time, we expected them not to overestimate their FOK performance because we predicted that weak levels of partial information was not likely to create an initial sense of high familiarity. Thus, their FOKs may be more accurate compared to the condition in which participants were given stimuli with strong partial information.

We expected participants to give comparable FOK judgments in response to either one of the stimuli sets as long as they respond under self-pace conditions. This was expected because they had enough time to evaluate the partial information produced. Thus, this evaluation process could result in more accurate FOK judgments compared to restricted time conditions, regardless of the amount of partial information retrieved. It was also possible that stimuli with weak partial information create feeling of not knowing, and this would result in a decrease in FOK magnitude, and possibly a decrease in FOK accuracy.

CHAPTER 2

EXPERIMENT 1

2.1 Method

2.1.1 Participants

A total of 97 undergraduate students (75 female) from Boğaziçi University who were enrolled in the Introduction to Psychology course participated in this experiment. Participants were given one point to be added to their overall course grade as compensation for participation.

2.1.2 Materials

In the encoding phase of this experiment, we used cue-target word pairs. All words in this experiment were chosen from the Turkish Word Norms (Tekcan & Göz, 2005). This database contains free association norms along with written word frequency, concreteness, and imagery ratings of 600 Turkish words chosen from the Words Frequency Dictionary of Written Turkish (Göz, 2003) based on two criteria. Among these 600 words, there is an equal number of high frequency, medium frequency and low frequency words (for both $N = 200$). Also, half of the words are concrete ($N = 305$; 51%), and the remaining half are abstract ($N = 295$; 49%). In order to manipulate levels of partial information in the encoding phase, we created two sets of target words that differed in association set sizes. The rationale for the selection of words was as follows:

We assumed that words that would lead to strong partial information would have a small association set and yet strong and distinct primary associations. In other words, the words in the strong partial information condition were words that may be associated with a small set of words, of which one dominated most responses. Thus, any partial information coming from the activation of such words would reflect the activation of nodes close to each other in an imaginary web of words and consequently this process would allow for the build-up of related information. Based on this rationale, words chosen for the strong partial information condition had 16 or fewer associations and the first association was identified by 35 or more people (out of a 100). For example, *bulut* (cloud) had a total of 14 associations and 55 out of 100 raters associating *yağmur* (rain) with this word.

For the words in the weak partial information condition, we chose words that had larger association sets and weaker primary associations. These words were typically associated with more idiosyncratic associations. Thus, any partial information coming from activation of such words may be less related with one another and possibly with weaker connections. To select these words, we chose words that had at least 39 associations and with 12 or less people associating to the same word. For example, *patron* (boss) had 48 associations and only 12 people had associated the same word, *iş* (job) with it.

After the selection of words, each category was examined for any possible outliers in terms of frequency, concreteness and imagery scores. Two outliers were found in the weak partial information group in terms of frequency and were excluded from the list. As a result, both weak partial information and strong partial information lists ended up having 27 words each (Appendix A).

Previous literature suggests that there may be concreteness and imagery differences between the two lists that we have generated (Galbraith & Underwood, 1973; Marschark & Hunt, 1989). The relational distinctiveness framework (e.g. Marschark & Hunt, 1989) states that lists consisting of words with strong primary associate and low number of associations (i.e. the strong partial info list) tend to be more concrete due to having item specific distinctiveness. On the other hand, the contextual variety hypothesis (e.g. Galbraith & Underwood, 1973), states that word lists consisting of words associated with a weak primary and high overall number of associations to be more abstract given the larger associative set. In our first group of words, strong partial information word group was higher in both imagery and concreteness scores; however, because we wanted to establish effects of partial information on retrieval and FOK, we intentionally tried to eliminate any differences on concreteness/abstractness and imagery across the two lists. Thus we inspected the lists on these criteria. The final lists contained 24 words each. These lists did not differ in terms of frequency, t (46) = 1.911, p = .062[1], r = .27; concreteness, t (46) = 1.507, p = .139, r = .22 or imagery, t (46) = 1.216, p = .266, r = .18. See Table 1 for descriptives. These lists were associated with 24 cue words with medium frequency levels (frequencies ranging from 22 to 92) and medium levels of associations, all have association sets ranging from 25 to 30 words. See Appendix B. We aimed to be free of the impact of cue familiarity in terms of memory and metamemory measures we used same cues across two different groups of target words.

---

[1] Frequency rating of Weak group (*M* = 94.83; *SD* = 101.53) was higher than frequency ratings of Strong group (*M* = 50; *SD* = 53.87). Effect size for this analyses did not qualify for a significant effect Cohen's d = .02.

Table 1. Descriptives for frequency, imagery and concreteness scores of target words.

	Weak		Strong	
	M	*SD*	*M*	*SD*
Frequency	94.83333	101.5389	50	53.87102
Imagery	4.24625	1.346034	4.827083	1.324073
Concreteness	4.689167	1.951423	5.236667	1.366457

We chose an additional 54 words for the recognition phase of the experiment. In the recognition phase, there were four alternative responses. One of these alternatives was the actual target word, and the other one was another target word from the encoding phase (paired with a different cue). The remaining two alternatives were either the word most frequently associated with the actual target (Appendix A) or a totally novel lure. These lures were of medium frequency (ranged from 22 to 98) and had mid-size sets of association (ranged from 25 to 30). See Appendix C.

2.1.3 Procedure

All phases of the experiment were administered via E-Prime 2 Professional Software (Psychology Software Tools, Pittsburgh, PA); the filler task, which consisted of an online survey, was administered via psychsurveys.org. Participants were tested individually. At the start of the experiment, we asked each participant to sign a consent form. The experiment consisted of encoding, filler task, cued recall, FOK judgment, recognition and confidence judgment phases. In the encoding phase, participants were randomly assigned to either weak or strong partial information condition and were asked to study the respective lists. After the encoding phase, participants engaged in a filler task, which took approximately 20 minutes. Then, they completed the cued recall task followed by the FOK judgment phase. In the cued recall phase, participants were presented with the cue words one by one and asked to recall the target of the cue that is presented in the encoding phase. After all the trials were completed, participants were asked for their FOK judgments. For all pairs, they were instructed to report their likelihood of recognizing the correct target amongst four choices. FOK ratings ranged from 0 (definitely cannot recognize) to 100 (definitely can recognize). The FOK judgment phase had two alternative versions; half of the participants received the restricted FOK judgment phase, in which participants were forced to answer the question in 2000ms. In the other condition, participants gave FOK responses at their own-pace. Although participants were given explicit instructions to respond as soon as possible in restricted time condition, participants in the self-paced time conditions were not instructed explicitly. The reason why they were not instructed to take as much time as they needed was to not prompt particular strategies. Once FOK judgments were given, participants moved onto the recognition phase and were asked to choose the target from four of the alternative options by clicking with their mouse on the correct target. After each recognition judgment, participants also gave confidence ratings for their recognition judgments that ranged from 0 (definitely not sure) to 100 (definitely sure).

2.2 Results

For each participant we calculated memory accuracy (recall and recognition), mean FOK strength and FOK accuracy (gamma calculations). Data from 17 participants were excluded because they either had zero recall accuracy or failed to provide data in the FOK judgment phase in the restricted time condition. Data from the remaining 80 participants were analyzed. Of the 80 participants, data from 5 participants were removed from the analyses due to being outliers in terms of either FOK reaction time ($N = 1$), FOK mean ($N = 2$) or recognition accuracy ($N = 2$). For the remaining sample, we provide the descriptives for memory and metamemory measures in the Table 2.

2.2.1 FOK response time

In the experiment, in the restricted time condition, participants had to give their FOK response within 2000ms. As a manipulation check, we wanted to determine whether people took longer in the self-pace condition. Results of the independent samples t-test revealed that in the self-pace condition ($M = 3539$; $SD = 934$), participants gave FOK responses slower than in the restricted time condition ($M = 1486$; $SD = 154$, $t\ (36) = 10.019$, $p = 001$, $r = .86$).

2.2.2 Memory performance

2.2.2.1 Recall

In order to compare the effect of word type on cued recall accuracy an independent sample t-test was conducted. Recall was higher in the weak partial information condition ($M = .42$; $SD = .20$) compared to strong condition ($M = .33$; $SD = .19$), $t\ (73) = 2.010$, $p = .048$, $r = .28$.

Table 2. Descriptives for Memory and Metamemory Measures

	Restricted		Self-paced		Main Effects								
	Weak	Strong	Weak	Strong	Timing			Word Type			Interaction		
	N = 20	*N* = 18	*N* = 18	*N* = 19									
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*	*F*	*P*	η^2_p	*F-t*	*p*	η^2_p	*F*	*p*	η^2_p
Reaction Time (ms)	1441 (168)	1536 (122)	3521 (1041)	3556 (849)									
Memory measures													
Recall (%)	.43 (.21)	.35 (.20)	0.41 (.21)	.31 (.18)	-	-	-	2.010	0.048*		-	-	-
Recognition (%)	.76 (.16)	.81 (.16)	.79 (.13)	.84 (.11)	0.936	0.337	0.013	2.252	0.138	0.031	0.007	0.935	0.000
Conditional recognition (%)	.35 (.10)	.48 (.16)	.40 (.12)	.53 (.12)	2.718	0.104	0.037	19.658	0.000*	0.217	0.009	0.924	0.000
Metamemory measures													
FOK Strength	68.70 (18.82)	51.69 (16.79)	61.44 (20.44)	61.90 (16.30)	0.123	0.726	0.002	3.890	0.052	0.052	4.327	0.041*	0.057
Correctly Recalled	86.26 (21.94)	71.82 (25.20)	91.36 (11.78)	97.09 (5.81)	13.381	0.000*	0.159	1.099	0.289	0.015	5.907	0.018*	0.077
Incorrectly Recalled	44.11 (21.46)	31.16 (20.85)	44.96 (20.66)	45.87 (15.67)	2.892	0.093	0.039	1.730	0.193	0.024	2.291	0.135	0.031
Correctly Recognized	71.98 (18.27)	55.40 (17.98)	67.77 (17.56)	66.24 (14.19)	0.685	0.411	0.010	5.118	0.027*	0.067	3.529	0.064	0.047
Incorrectly Recognized	32.56 (18.73)	19.10 (17.07)	32.25 (25.89)	31.36 (24.19)	1.413	0.239	0.020	2.035	0.158	0.028	1.561	0.216	0.022
FOK accuracy (Gamma)	.37 (.64)	.39 (.60)	.66 (.23)	.59 (.33)	4.789	0.032*	0.063	0.049	0.825	0.001	0.168	0.683	0.002

Notes.
Significant findings are indicated with *
The *t* score within the word type main effects is only for the recall accuracy. We did not compare the effect of timing for recall accuracy because the timing manipulation was subsequent to recall.

2.2.2.2 Recognition

In order to compare the effect of word type as well as FOK timing on recognition accuracy, we conducted a 2 (Word Type: Strong vs. Weak) X 2 (Timing: Restricted vs. Self-paced) Between Subjects ANOVA. These analyses did not reveal any significant findings (see Table 2). Participants' recognition performance was not affected by either of the manipulations. However, when we calculated the conditional recognition performance (correct recognition of non-recalled items in the cued recall phase) of the participants we found a main effect of word type, F (1, 74) = 19.658, MSE = .016, p = .001, η^2_p = .217. Participants in the strong partial information condition had better accuracy (M = .50; SD = .14) compared to participants given weak partial information (M = .37; SD = .11).

2.2.3 Metamemory performance

2.2.3.1 FOK magnitude

We predicted that both timing and word type would have an effect on FOK judgment magnitude. In particular, we predicted that participants in the self-paced condition would give FOK judgments of similar magnitude for strong and weak words. On the other hand, participants who were given strong words were expected to give highest FOK judgment in the restricted time condition. In order to test this, we conducted a 2 (Word Type: Strong vs. Weak) X 2 (Timing: Restricted vs. Self-paced) Between Subjects ANOVA on mean FOK ratings. These analyses revealed a marginal effect of word type, F (1, 74) = 3.890, MSE = 329, p = .052, η^2_p = .052. Participants given the weak word (large association set) list gave higher FOK ratings (M = 65.07; SD = 19.68) compared to participants who were given the strong word (small association set) (*M* = 56.79; *SD* = 17.11). There was no effect of timing, $p = .726$. There was an interaction, $F(1, 74) = 4.327$, $MSE = 329$, $p = .041$, $\eta^2_p = .057$. As we predicted there was no difference in performance in the self-pace condition for weak and strong partial information lists. However, participants in the restricted time condition differed, $t(36) = 2.926$, $p = .006$, $r = .44$. Contrary to our expectation, participants gave higher FOK ratings for weak words (*M* = 68.70; *SD* = 18.82) in the restricted time condition, compared to strong words (*M* = 51.69; *SD* = 16.79). See Figure 1.

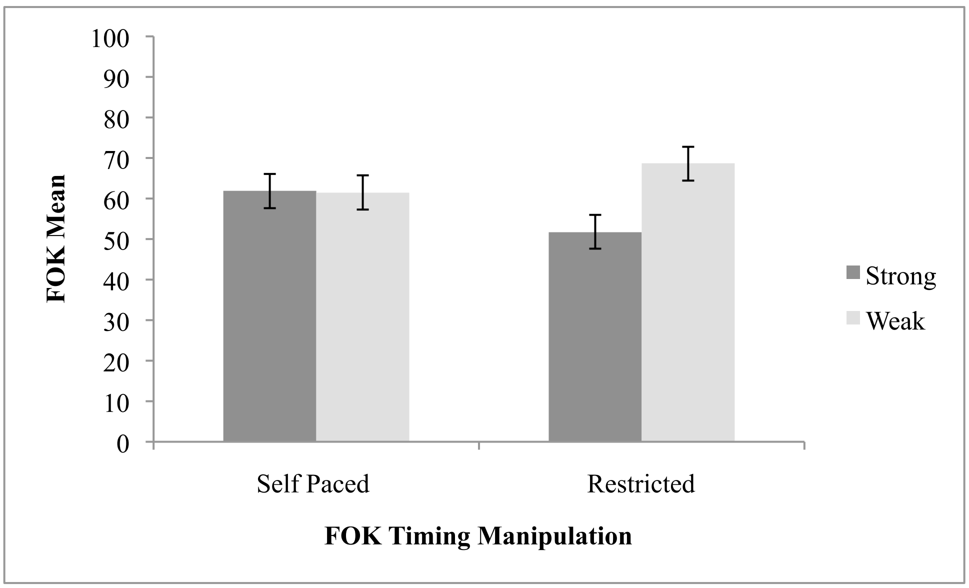

Figure 1. FOK magnitude and standard error of the means as a function of timing and word type

2.2.3.2 FOK accuracy

We predicted that both timing and word type would have an effect on FOK accuracy. In particular, we expected that participants in the self-paced time conditions would not differ across word type; however, participants in the restricted time conditions would have lowest levels of FOK accuracy for the strong words. In order to test our expectations, we calculated gamma scores for FOK accuracy. Results of the 2 (Word Type: Strong vs. Weak) X 2 (Timing: Restricted vs. Self-paced) Between Subjects ANOVA revealed an effect of timing, F (1, 74) = 4.789, MSE = .235, p = .032, η^2_p = .063. Participants who gave self-paced responses (M = .62; SD = .28) were more accurate than participants in the restricted time condition (M = .38; SD = .61). There was no effect of word type and no interaction, ps > .6. See Figure 2.

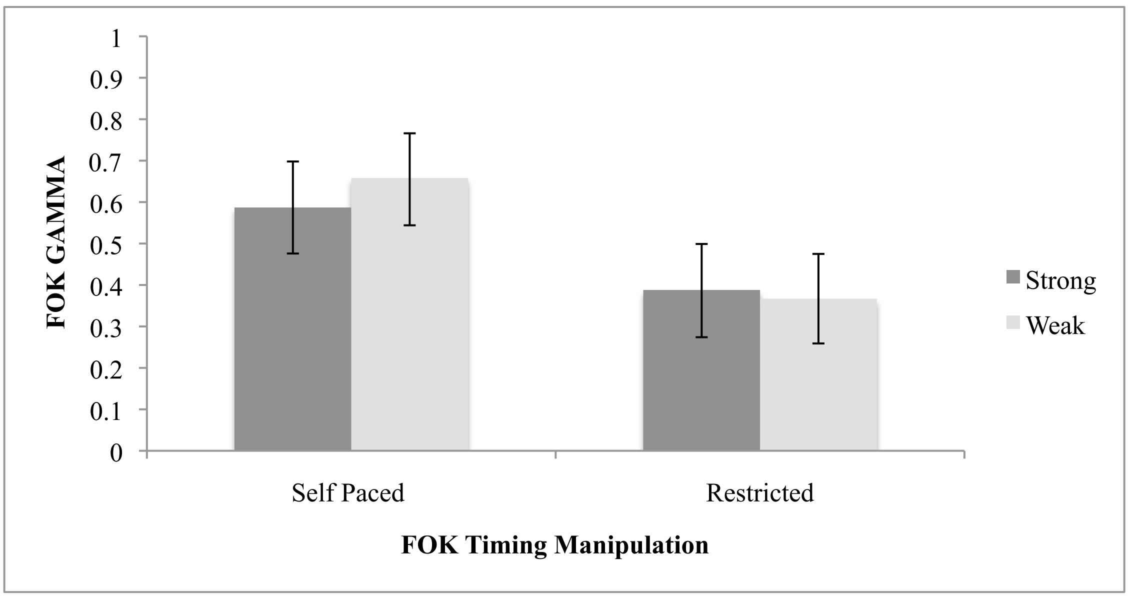

Figure 2. FOK accuracy and standard error of the means as a function of timing and word type

2.3 Discussion

In Experiment 1, we found the effect of timing only for FOK accuracy, in which participant who responded in self-paced time were more accurate compared to restricted time conditions. There was no effect of word type or interaction for this measure. This result could be interpreted as such: people who were given enough time to think while making their FOK judgments had more time to evaluate their retrieved information, compared to participants who were given restricted time. This result was observed regardless of the word type. It is evident from the previous literature that people respond differently when they are given restricted time compared to when they respond at their own pace. Following Koriat and Levy-Sadot (2001), such a difference might result from the total amount of retrieved partial information. Given enough time, people remember more information related to the target and they also have more time to evaluate that information. Therefore, people make higher FOK judgments, which, in turn correlate with the likelihood of recognition (see also Benjamin, 2005). Similarly, Thomas and her colleagues (2011) showed that forcing people to reply in a restricted time window during the FOK stage blocks evaluative processing, and thus reduce FOK accuracy.

We found an effect of word type on FOK magnitude. Contrary to our expectations, we found that recall and FOK judgments were higher for the weak partial information group. At the development of this experiment, the weak words had been conceptualized as weaker in terms of their level of association because they were more likely to have idiosyncratic associations, which did not cluster around a certain concept in a given web of words. Thus, we would have expected them to get lower FOK judgments. Given the results, we suspected that higher recall performance on the weak partial information condition may be driven by the possible relationships between the cue-target word pairs. Although we had inspected possible imagery, concreteness and frequency differences between target words across the two lists, we did not actually control for how strongly the cue and the target could be imagined together. Therefore, we conducted a follow-up study to test for the possible effect of distinctiveness, imagery and concreteness scores of the cue-target word pairs. One hundred eight students taking the introduction to psychology course rated cue-target word pairs on distinctiveness within all other cue-target word pairs in the experiment, along with imagery and concreteness scores. They first indicated the number of syllables of each cue-target word pairs. This manipulation was done to ensure that each participant saw all the word pairs before giving ratings for the distinctiveness, imagery, and concreteness scores. After the syllable task, participants rated distinctiveness scores of the all cue-target pairs. Participants instructed to give highest ratings (7) to the most distinctive word pairs from the list and lowest ratings (1) to the least distinctive word pairs, for the pairs, which did not belong to these categories they were instructed to choose adequate rating ranging from 2 to 6. All the word pairs were shown in one screen and in a randomized format for each participant. After all distinctiveness ratings were completed, all participants rated imagery and concreteness of the word pairs, and they were randomly assigned to either imagery-first or concreteness-first conditions. Similar to distinctiveness ratings, all word pairs were shown in one screen and randomized for each participant. Although the words pairs did not differ in terms of distinctiveness and concreteness across the lists (ts < 1.14, ps $> .1$), we found a marginal effect of imagery, t (46) = 1.139, $p = .058$, $r = .17$. Cue-target pairs with weak partial information targets were rated higher in imagery (M = 4.27; SD = .78) compared to cue-target word pairs with strong partial information targets (M = 3.81; SD = .85).

Based on the follow up experiment, we anticipated that the higher recall and mean FOK magnitude for the weak group might have resulted from the higher imagery of the cue-target pairs. It is evident from the previous literature that imagery had an effect on both memory and metamemory performance (see Hertzog et al., 2010; 2014; Thomas et al., 2012). Therefore, in Experiment 2 we wanted to eliminate this possible confound.

CHAPTER 3

EXPERIMENT 2

In Experiment 2, in order to eliminate the effect of higher imagery between cue and target in weak partial information condition, we rearranged the cue-target pairs that were rated higher in the both weak and strong partial information conditions. It should be noted that both cue and target words were identical to those used in Experiment 1. The only difference was that some of the target words were paired with other cues from the set to create similar levels of cue-target pair imagery across the two lists.

3.1 Method

3.1.1 Participants

A total of 106 undergraduate students (73 female) from Boğaziçi University, who were enrolled in the 'Introduction to Psychology' course, participated in this experiment for one course credit for compensation.

3.1.2 Materials

Encoding, cued recall, FOK judgment, recognition and confidence judgment tasks were prepared and presented by E-Prime 2 Professional Software (Psychology Software Tools, Pittsburgh, PA). Filler task and survey for the cue-target word pairs was prepared and presented by psychsurveys.org.

3.1.3 Procedure

Procedure of the Experiment 2 was identical to Experiment 1, except, in Experiment 2 participants rated the distinctiveness, imagery and concreteness of each cue-target word pairs at the end of the experiment.

3.2 Results

There were no differences between the two lists in terms of cue-target pairs on distinctiveness, imagery and concreteness, all *p*s >.15. For each participant we calculated memory accuracy (recall and recognition), mean FOK magnitude, and FOK accuracy (gamma) scores. Data from 17 participants were eliminated because they either had zero recall accuracy or failed to provide data in the FOK judgment phase in the restricted time condition. Data from the remaining 89 participants were analyzed. From 89 participants, a total of nine participants were removed from the analyses because they were outliers in terms of FOK reaction time ($N = 4$), FOK mean ($N = 2$), FOK gamma ($N = 2$) and recognition accuracy ($N = 1$). For the remaining sample of 80, descriptives for memory and metamemory measures are summarized in the Table 3.

3.2.1 Reaction time

As in Experiment 1, we wanted to confirm that people took longer in the self-pace compared to the restricted time condition while giving their FOK judgments. This was indeed the case, participants responded slower in the self-pace ($M = 3815$; $SD = 1019$) than in the restricted ($M = 1519$; $SD = 133$) condition, $t\,(73) = 1.817$, $p = 001$, $r = .21$.

3.2.2 Memory performance

3.2.2.1 Recall and recognition

We found no differences in the cued recall accuracy and recognition accuracy; *p*s > .1 (see Table 3). When we calculated the conditional recognition scores (recognition of non-recalled items in the cued recall phase) and conducted a 2 (Word Type: Strong vs. Weak) X 2 (Timing: Restricted vs. Self-paced) Between Subjects ANOVA, we found an interaction, F (1, 76) = 5.184, *MSE* = .017, p = .026, η^2_p = .064. Post-hoc tests revealed that there was no difference in the self-paced condition (p = .535), however, participants in the restricted time condition were more accurate in the strong partial information condition compared to the weak partial information condition, t (42) = 2.661, p = .011, r = .38.

In order to investigate the recognition performance further, we coded the recognition errors of the participants. We specifically coded whether they misrecognized the first association, another target from the list or the lure instead of the actual target. Participants showed comparable performance for the first association and other target choices. However, participants who were in the strong partial information condition showed a significantly decreased preference to lures (4 %) compared to participants who were in weak condition (7 %), t (78) = 2.016, p = .047, r = .28.

Table 3. Descriptives for Memory and Metamemory Measures

	Restricted		Self-paced		Main Effects									
	Weak	Strong	Weak	Strong	Timing			Word Type			Interaction			*MSE*
	N = 25	*N* = 19	*N* = 15	*N* = 21										
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*	*F*	*p*	η^2_p	*F - t*	*p*	η^2_p	*F*	*p*	η^2_p	
Reaction Time (ms)	1511 (133)	1530 (136)	3744 (773)	3868 (1179)										
Memory measures														
Recall (%)	0.33 (0.21)	0.26 (0.19)	0.26 (0.12)	0.34 (0.21)	-	-	-	0.093	0.926	-	-	-	-	-
Recognition (%)	0.76 (.17)	0.8 (0.12)	0.77 (0.14)	0.83 (0.15)	0.403	0.527	0.005	1.886	0.174	0.024	0.115	0.736	0.002	0.022
Conditional recognition (%)	0.43 (.12)	0.54 (0.15)	0.52 (0.11)	0.49 (0.13)	0.343	0.56	0.004	1.895	0.173	0.024	5.184	0.026*	0.064	0.017
Metamemory measures														
FOK Mean	59.55 (18.83)	49.8 (22.49)	50.87 (12.02)	65.62 (20.46)	0.667	0.417	0.009	0.328	0.569	0.004	7.876	0.006*	0.094	386
Correctly Recalled	83.05 (20.18)	76.55 (29.36)	94.51 (4.95)	94.96 (10.61)	11.754	0.001*	0.134	0.483	0.489	0.006	0.636	0.428	0.008	366
Incorrectly Recalled	42.04 (18.39)	32.88 (17.1)	35.4 (13.05)	52.13 (22.4)	2.267	0.136	0.029	0.815	0.37	0.011	9.541	0.003*	0.112	339
Correctly Recognized	65.32 (18.38)	52.05 (22.83)	56.76 (10.89)	68.79 (18.08)	0.959	0.331	0.012	0.022	0.882	0	9.157	0.003*	0.108	337
Incorrectly Recognized	26.38 (19)	21.06 (15.99)	31.17 (18.17)	44.49 (33.09)	7.346	0.008*	0.088	0.591	0.444	0.008	3.21	0.077	0.041	523
FOK Gamma	0.35 (0.66)	0.28 (0.57)	0.47 (0.33)	0.2 (0.72)	0.029	0.865	0	1.496	0.225	0.019	0.508	0.478	0.007	0.373
FOK Gamma Inc Recog	0.22 (0.7)	0.14 (0.6)	0.26 (0.38)	0.03 (0.76)	0.057	0.813	0.001	1.053	0.308	0.014	0.27	0.605	0.004	0.419

Notes.
Significant findings are indicated with *
The *t* score within the word type main effects is only for the recall accuracy. We did not compare the effect of timing for recall accuracy because timing manipulation was subsequent to recall.

3.2.3 Metamemory performance

3.2.3.1 FOK magnitude

As in Experiment 1, we predicted that both timing and word type would have an effect on FOK magnitude. In particular, we predicted that in the self-paced condition FOK magnitude would not differ for the strong and weak lists. Moreover, participants who were given the strong association list were expected to give the highest FOK judgment in the restricted time condition. In order to test this, we conducted a 2 (Word Type: Strong vs. Weak) X 2 (Timing: Restricted vs. Self-paced) Between Subjects ANOVA on mean FOK ratings. Results revealed that there was no effect of word type or timing, *p*s > .5 (see Table 3). There was an interaction but not in the expected direction, F (1, 76) = 7.876, *MSE* = 386, p = .006, η^2_p = .094. As can be seen in Figure 3, this interaction was driven by the equal levels of performance in restricted time condition (p = .125) in contrast to differences in the self-paced condition, t (34) = 2.495, p = .018, r = .34. In the self-paced condition, participants gave higher FOK ratings to strong (M = 65.62; SD = 20.46) than weak words (M = 50.87; SD = 12.02).

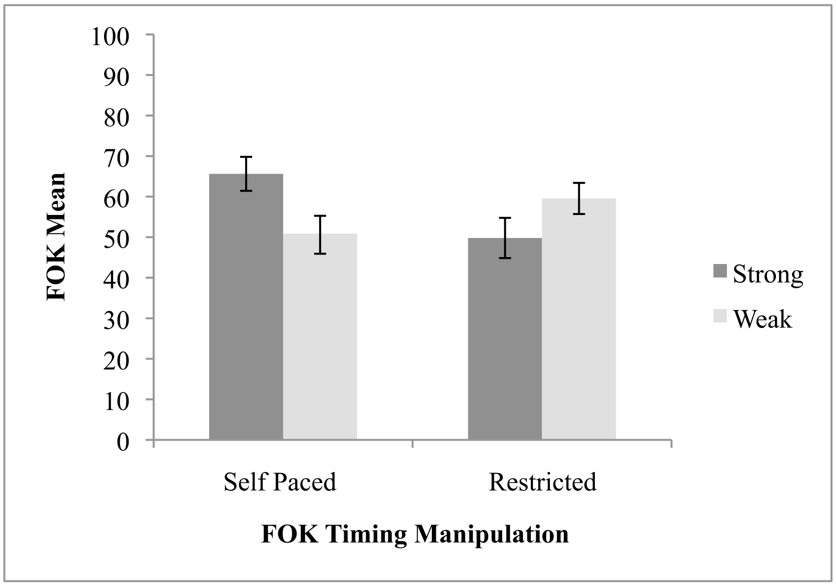

Figure 3. FOK magnitude and standard error of the means as a function of timing and word type

When participants are making FOK judgments it was possible that they are affected by their earlier recall judgments. Thus, in order to investigate for possible differences, we conducted our analyses on correctly recalled and incorrectly recalled trials. First of all, we ran a 2 (Word Type: Strong vs. Weak) X 2 (Timing: Restricted vs. Self-paced) Between Subjects ANOVA on mean FOK judgments for correctly recalled trials. Participants in the self-paced condition (*M* = 94.75; *SD* = 8.72) gave higher ratings compared to participants in the restricted time conditions (*M* = 79.80; *SD* = 17.56), F (1, 76) = 11.754, *MSE* = 483, p = .001, η^2_p = .134. There was no effect of word type and no interaction; *p*s > .4 (see Table 3). As can be seen in figure 4, post-hoc analyses revealed that effect of timing was observed for both strong words (p = .011) and weak words (p = .038). See Figure 4.

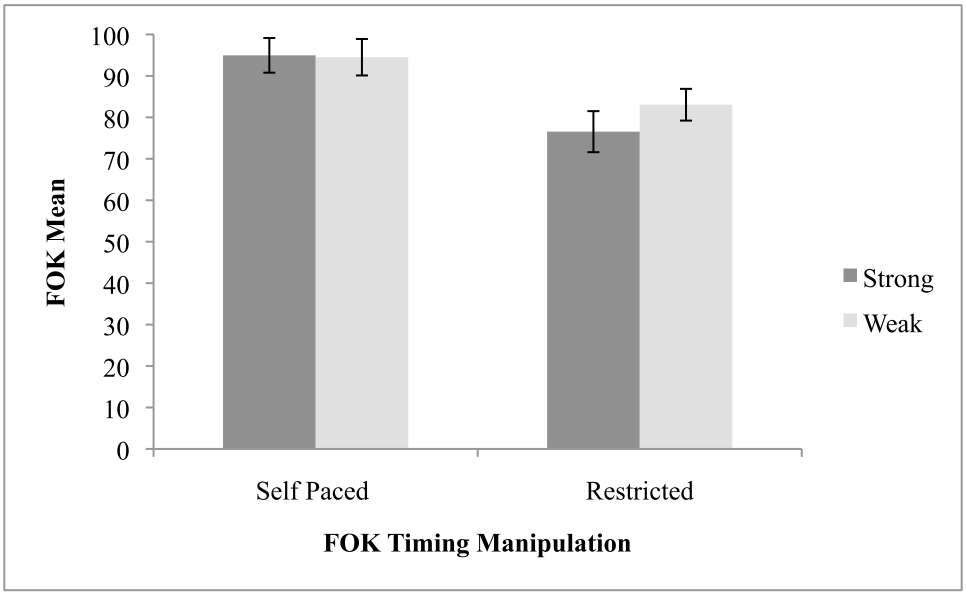

Figure 4. FOK magnitude and standard error of the means as a function of timing and word type for correctly recalled trials

Secondly, results of the 2 (Word Type: Strong vs. Weak) X 2 (Timing: Restricted vs. Self-paced) Between Subjects ANOVA on mean FOK ratings for incorrectly recalled trials revealed that there was no effect of timing and word type *p*s > .1 (see Table 3). There was an interaction, F (1, 76) = 9.541, *MSE* = 339, p = .003, η^2_p = .112. This interaction was driven by the equal levels of performance in restricted time condition (p = .099). However, participants in the self-paced time condition differed, t (34) = 2.589, p = .014, = .41. In the self-paced condition participants gave higher FOK ratings for words with strong partial information (M = 52.13; SD = 22.40), compared to weak partial information (M = 35.40; SD = 13.05), t (38) = 2.584, p = .014, r = .39. These results mimic the findings for the overall FOK magnitude (See the similarity between Figure 3 & 5).

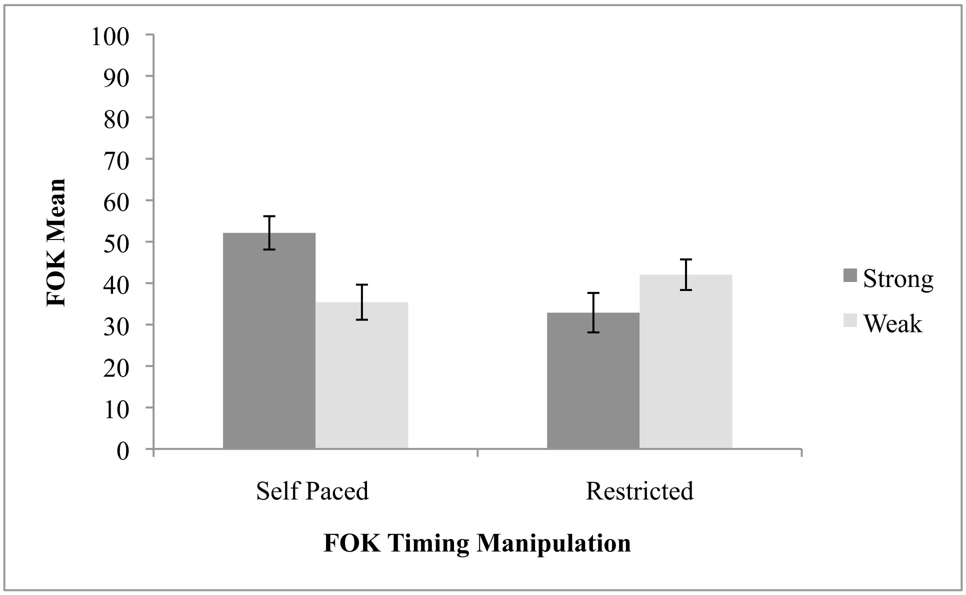

Figure 5. FOK magnitude and standard error of the means as a function of timing and word type for incorrectly recalled trials

Along with the recall accuracy we predicted that recognition accuracy could be an important discriminator for making FOK judgments. Although FOK judgment phase always precedes recognition, participants made their FOK judgments as a prediction for their upcoming recognition performance. Thus, we conducted analyses on both correctly recognized and incorrectly recognized trials. First of all, results of the 2 (Word Type: Strong vs. Weak) X 2 (Timing: Restricted vs. Self-paced) Between Subjects ANOVA on mean FOK ratings for correctly recognized trials revealed that there was no effect of timing and word type *p*s > .3 (see Table 3). As can be seen in Figure 6, there was a crossover interaction, F (1, 76) = 9.157, *MSE* = 337, p = .003, η^2_p = .108. Participants in the restricted time condition performed differently across word types, t (42) = 2.137, p = .038. Specifically, words with strong partial information (*M* = 52.05; *SD* = 22.83) were given lower FOK ratings than weak (*M* = 65.32; *SD* = 18.38). Similarly, participants in the self-paced time condition differed, t (34) = 2.292, p = .028, r = .31). As can be seen in Figure 6, in the self-pace condition, strong words (M = 68.79; SD = 18.08) were rated higher compared to weak words (M = 56.76; SD = 10.89). Moreover, for strong words, participants gave higher ratings in the self-pace condition, t (38) = 2.584, p = .014, r = .39.

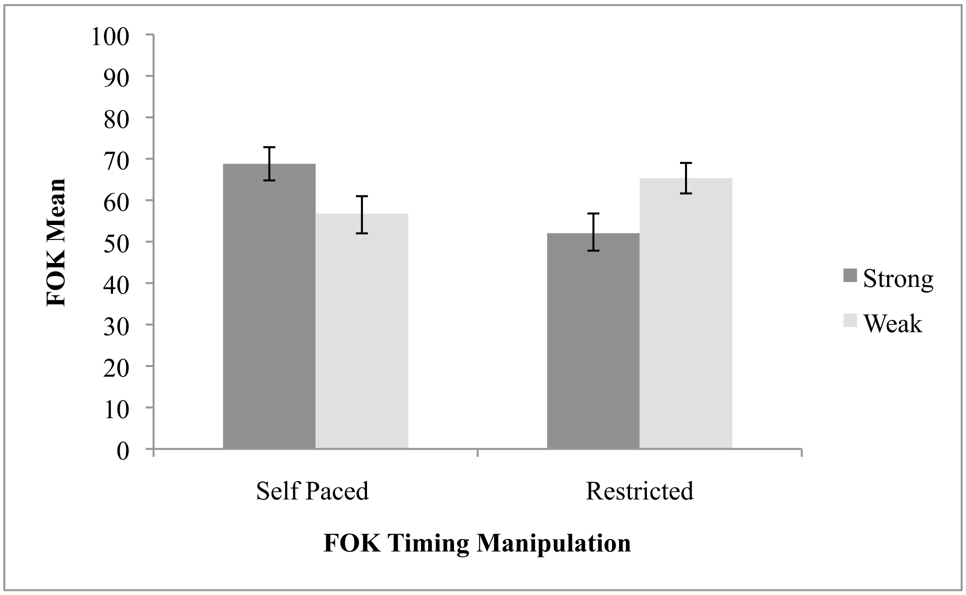

Figure 6. FOK magnitude and standard error of the means as a function of timing and word type for correctly recognized trials

Lastly, 2 (Word Type: Strong vs. Weak) X 2 (Timing: Restricted vs. Self-paced) Between Subjects ANOVA on mean FOK ratings for incorrectly recognized trials revealed that there was an effect of timing, F (1, 76) = 7.346, MSE = 523, p = .008, η^2_p = .088. As can be seen from the Figure 7, participants in the self-paced condition (M = 37.83; SD = 3.87) gave higher ratings compared to participants in the restricted time condition (M = 23.72; SD = 3.48). There was no effect of word type and no interaction,

p > .07 (see Table 3). Given that incorrectly recognized trials less than 24% of the data, we evaluated these results with skepticism.

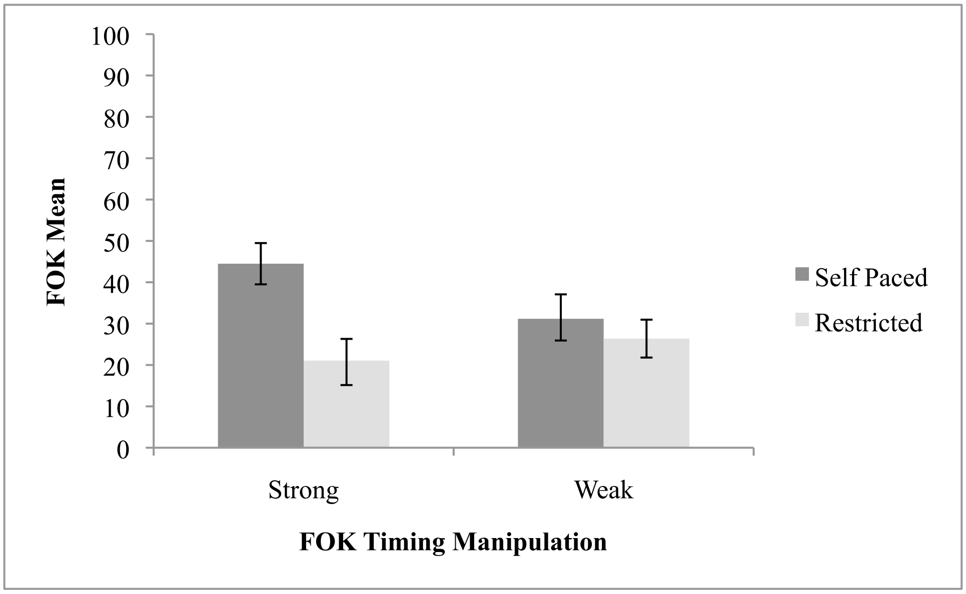

Figure 7. FOK magnitude and standard error of the means as a function of timing and word type for incorrectly recognized trials

As there was a wider range in FOK RT in the self-paced condition than the restricted time condition, we wanted to investigate whether within this group there were any differences in FOK ratings based on RT. First of all, we divided the participants into two groups: fast responders and slow responders, through a median split of their FOK reaction times. We looked at whether fast responders performed differently compared to slow responders. Results of the 2 (Word Type: Strong vs. Weak) X 2 (Timing: Restricted vs. Self-paced) Between Subjects ANOVA on mean FOK ratings did not revealed any group differences (all ps >.09). We also compared FOK ratings for the fastest (the 25^{th} percentile or lower) and the slowest (75^{th} percentile or higher) FOK trials, again based on the FOK response time. Similarly, results of the 2 (Word Type:

Strong vs. Weak) X 2 (Timing: Restricted vs. Self-paced) Between Subjects ANOVA did not revealed any significant findings (all *p*s > .20). Additionally, we furthered our investigation in a trial basis. We detected 5 participants who responded faster (<2000ms) in more than half of the trials. We compared their mean and median FOK magnitude ratings across slowest and fastest trials; yet, we could not find any significant pattern.

3.2.3.2 FOK accuracy

We predicted that both timing and word type would have an effect on the accuracy of FOK judgments. In particular, we expected that participants in the self-paced time conditions would not differ with respect to accuracy across word lists. However, we predicted that participants in the restricted time conditions would have lowest levels of accuracy for the strong words. Comparison of gamma scores for FOK accuracy no main effects or interaction. Specifically the results of the 2 (Word Type: Strong vs. Weak) X 2 (Timing: Restricted vs. Self-paced) Between Subjects ANOVA on FOK gamma calculations revealed no effect of FOK response timing, word type and no interaction *p*s > .2 (see Table 3).

3.3 Discussion

The main purpose of the Experiment 2 was to control for the effect of imagery in the cue-target word pairs between two word groups. Results showed that rearrangement of cue – target word pairs eliminated the effect of imagery, and thus lists were more comparable, and this manipulation changed the pattern of results between Experiment 1 & 2.

Although there was an effect of timing for FOK accuracy in Experiment 1, such that in self-paced conditions people had higher FOK accuracy contrary to our predictions, this effect was not present in Experiment 2. This finding – i.e. the lack of effect on FOK accuracy was comparable with the literature. Although many studies show differences in the FOK magnitude ratings across different manipulations, often these effects were not observed for FOK accuracy (see Schwartz et al., 2014; Thomas et al. 2012). These manipulations include using stimuli that inherited different levels of partial information (Thomas et al., 2012), as in this thesis, or experimentally manipulating the number of information given to the participants (Schwartz et al., 2014). Moreover, in the present experiment, higher FOK judgments were given in the self-paced conditions especially for the strong partial information words. The lack of an effect on FOK accuracy could be explained through Koriat's (1993) accessibility model, in which, more partial information does not necessitate the retrieval of correct partial information, which in turn, does not lead to higher accuracy. It is highly possible that participants in self-paced conditions gave higher FOK judgments to strong words acknowledging the incorrect partial information during retrieval.

As we predicted, there was an interaction of word type and timing on FOK magnitude. In particular, when people were given restricted time, their judgments did not differ across word type; however, when people were allowed to respond at their own pace, people gave higher FOK judgments for strong than weak partial information words. Further investigation of FOK strength in terms of recall accuracy revealed two different patterns. First of all, the pattern of FOK judgments for the incorrectly recalled trials perfectly mimicked the pattern of all FOK judgments. Both for incorrectly recalled trials and overall, when people were in the self-paced condition, people gave higher FOK ratings for strong partial information words than weak partial information words. Interestingly however, FOK judgments for the correctly recalled trials were only higher in the self-paced trials.

One possible interpretation for the higher FOK judgments in the strong self-paced condition may be related to the possible differences in strategy use in self-paced and restricted conditions (see Benjamin, 2005; Koriat & Levy-Sadot, 2001). In both of these experiments, researchers showed that participants respond according to the accessibility of the target when given enough time. Thus, in this particular experiment more time and higher level of partial information promoted higher FOK judgments in strong partial information and self-paced condition. This interpretation is also in line with Koriat's (1993) argument that amount of partial information increases as the level of partial information of the target at hand increases. Moreover, Thomas and her colleagues (2012) found that meaningful stimuli, such as nameable pictures drive higher FOK judgments. Our word type manipulation bears both quantitative (i.e. number of associations) and qualitative (i.e. association strength) differences between strong and weak words. Thus, higher judgments for the strong words could be interpreted as driving higher judgments due to inherent qualitative advantages. Differences in the findings for the correctly recalled trials may have to do with the accuracy of partial information provided. It is possible that people used their recall performance as a strong cue for the future recognition performance and gave higher FOK judgments for the correctly recalled trials in self-paced conditions regardless of the word type. Thomas et al. (2011) also found that the accuracy of retrieved partial information increased the magnitude of FOK judgments. Moreover, changes in the pattern of FOK judgments according to recall accuracy were reasonable. It was found that participants employed different strategies during FOK judgment phase when they are asked to provide FOK judgments for all trials or incorrect trials (Schwartz, Boduroglu, & Tekcan, 2015). This study illustrates that participants use recall as a cue for later recognition performance.

Moreover, we also calculated the FOK judgments for correctly recognized trials. The rationale for this calculation rests in the fact that people made their FOK judgments in order to project their upcoming performance in the recognition phase of the experiment. Thus, FOK magnitude ratings especially for the correctly recognized trials could be informative, and these trials also constituted a higher percentage of the trials (> 76%). In the self-paced condition people gave higher ratings to strong words, however, in the restricted time condition people gave higher ratings to weak words. Moreover, within the strong words, self-paced condition, again, resulted in higher ratings. This finding represents the differences in strategy use in self-paced and restricted time conditions (see Benjamin, 2005; Koriat & Levy-Sadot, 2001). These results can also be interpreted as follows. According to the New Theory of Disuse, in the initial stages of remembering, words with high association set creates a stronger retrieval strength; however, words with stronger interassociation has stronger storage strength (Bjork & Benjamin, 2011; Bjork & Bjork, 1992). Given that FOK is being an inferential judgment related to retrieval, the new theory of disuse could be applied to FOK judgments and could explain the pattern of results for the correctly recognized trials. In our study, target words that had a weak first association, also had idiosyncratic but larger association sets. Thus they may have stronger retrieval strength due to a load of different cues at the same time and this may reveal itself when people gave FOK judgments in the condition when they do not have time to evaluate their response. However, when people are given enough time so they can judge the effect of the stronger association between two words, they may thus be able to evaluate storage strength and respond accordingly. In our study, target words that had strong first association but a small association set had stronger associations, thus given enough time people could have judged the level of interassociation for the target words and made their FOK judgments accordingly.

Overall, results of the study implements that restricting the timing of FOK judgment restricts the accumulation of quality of information. In this study, strong words bear higher levels of partial information in terms of qualitative differences compared to weak words. Thus, higher FOK magnitude judgments for the strong words in self-paced conditions could have resulted in the accumulation of more qualitative information with given time.

Both in Experiment 1 and Experiment 2 we experimentally manipulated the levels of partial information of the target words used in the study. We presented participants with either words with strong partial information or weak partial information. In a further experiment we wanted to check whether our manipulation was evident in participants' explicit responses.

CHAPTER 4

EXPERIMENT 3

In Experiment 3, we wanted to collect data on the partial information generated regarding the target words. There have been several attempts that asked participants to retrieve partial information following FOK judgments (e.g. Schwartz et. al, 2014; Thomas et al., 2011). For example, Schwartz and his colleagues (2014) presented participants with pictures of imaginary animals along information related to these animals (county, diet, weight). Later in the testing phase, they asked participants to retrieve information regarding these animals. They found increase in FOK magnitude with the increase in information provided regarding the animals. Additionally, Thomas and her colleagues (2011) found that quality of the retrieved information regarding the target influences FOK judgments for both older and younger adults. However, older adults require explicit retrieval of this information. In this study, apart from measuring the magnitude of partial information load for both weak and strong words, we wanted to investigate the quality of partial information and whether the quality has any significant relationship with FOK accuracy or FOK magnitude.

We introduced a new phase in between the FOK judgment and recognition phases in which participants were asked to provide retrieved partial information related to the given target words they had seen in the study phase.

4.1 Method

4.1.1 Participants

A total of 57 undergraduate students (42 female) from Boğaziçi University, who were enrolled in the ‘Introduction to Psychology’ course, participated in this experiment. Participants received one point course credit for compensation their participation.

4.1.2 Materials

Cue--target word pairs were identical to Experiment 2. Encoding, cued recall, FOK judgment, partial information, recognition and confidence judgment tasks were prepared and presented by E-Prime 2 Professional Software (Psychology Software Tools, Pittsburgh, PA). The filler task was prepared and presented by psychsurveys.org.

4.1.3 Procedure

Experiment 3 was identical to Experiment 2, except, in Experiment 3 after completing the FOK judgments, participants moved onto the partial information phase. In this phase, participants were again provided with the cue words one by one and asked to provide any information that came to their mind in relation to the target word that was presented along with the cue word. They were instructed that this information could be the first letter of the target word, number of syllables of the target word, any word or concept associated with the target word. Participants were encouraged to write all pieces of partial information as possible. After the partial information phase, participants completed the recognition and confidence judgment phases of the experiment.

4.2 Results

For each participant we calculated memory accuracy (recall and recognition), mean FOK strength and FOK accuracy (gamma) scores. A total of nine participants were removed from the analyses due to being outliers in terms of FOK reaction time ($N = 3$), FOK mean ($N = 2$), FOK gamma ($N = 2$), recognition accuracy ($N = 1$) and conditional recognition accuracy ($N = 1$). The descriptives for memory and metamemory measures for the remaining sample are summarized in the Table 4.

4.2.1 Reaction time

As in Experiment 1 and Experiment 2 we compared the FOK reaction times of self-paced and restricted time conditions. Participants took significantly longer in self-paced conditions ($M = 3361$; $SD = 922$) compared to restricted time conditions ($M = 1585$; $SD = 149$), $t\ (41) = 9.115$, $p = .001$, $r = .82$.

4.2.2 Memory performance

4.2.2.1 Recall and recognition

We found no differences for recall, recognition and conditional recognition scores of the participants across word type and timing manipulation, ps > .2 (see Table 4).

4.2.2 Metamemory performance

4.2.2.1 FOK magnitude

We predicted that both timing and word type would have an effect on FOK judgment magnitude and we expected results to be comparable with Experiment 2. However, as can be seen in Figures 8 and 9, results of a 2 (Word Type: Strong vs. Weak) X 2 (Timing: Restricted vs. Self-paced) Between Subjects ANOVA on all mean FOK ratings and mean FOK ratings of incorrectly recalled trials revealed no significant effects, *p*s > .1 (see Table 4).

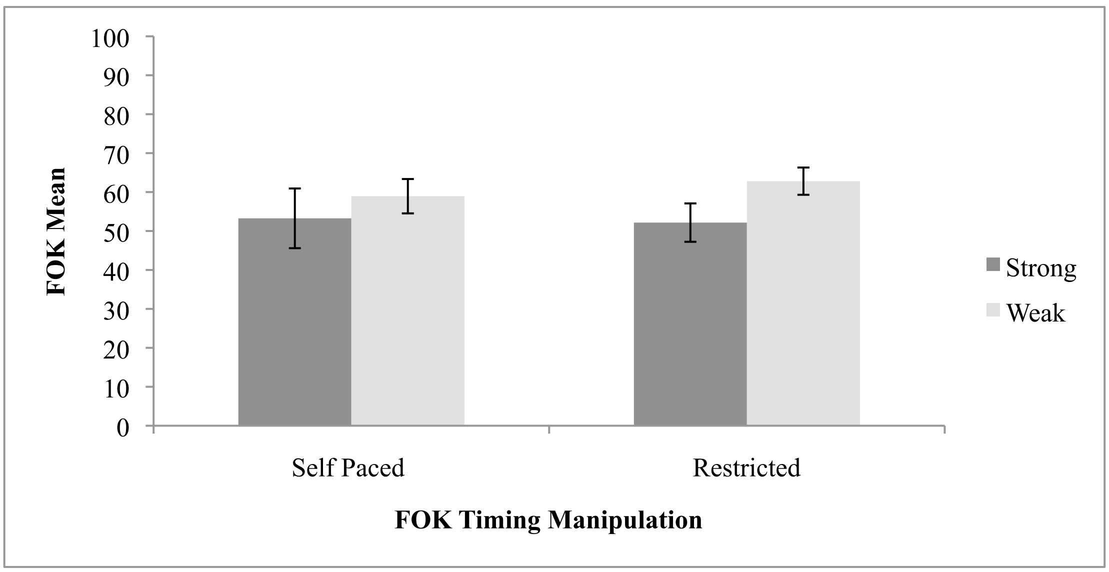

Figure 8. FOK magnitude and standard error of the means as a function of timing and word type for all trials

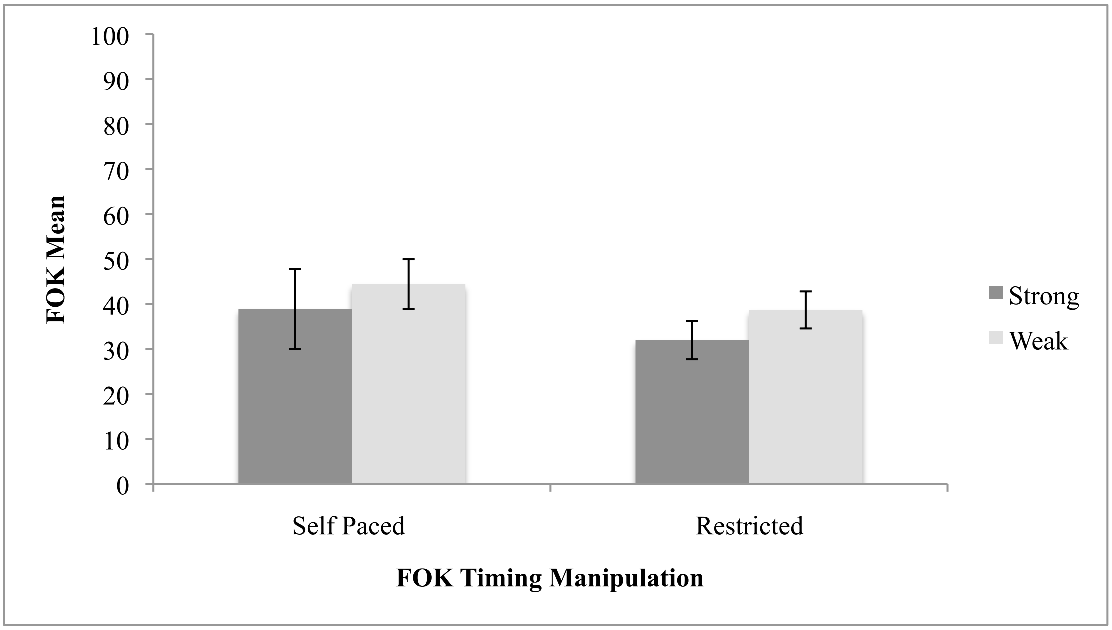

Figure 9. FOK magnitude and standard error of the means as a function of timing and word type for incorrectly recalled trials

As in Experiment 2, when we looked at the correctly recalled trials, FOK ratings were higher in the self-paced (*M* = 96.41; *SD* = 6.99) than in the restricted time conditions (*M* = 84.12; *SD* = 15.63), *F* (1, 39) = 9.841, *MSE* = 148, *p* = .003, η^2_p = .020. Figure 10 depicts FOK magnitude and standard error of the means as a function of timing and word type for correctly recalled trials.

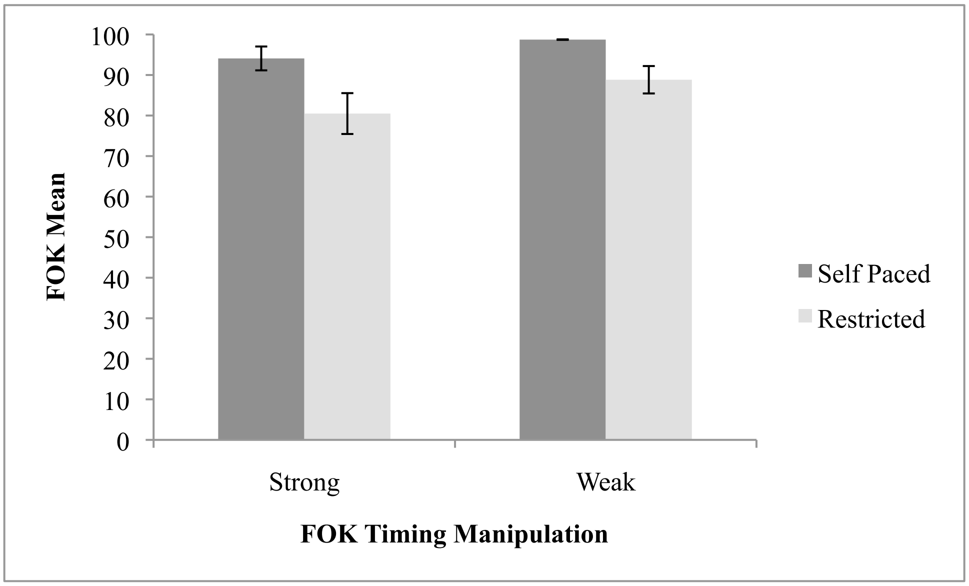

Figure 10. FOK magnitude and standard error of the means as a function of timing and word type for correctly recalled trials

Similar to Experiment 2, there was a wider range in FOK RT in the self-paced condition than the restricted time condition, we wanted to investigate whether within this group there were any differences in FOK ratings based on RT. First of all, we divided the participants into two groups: fast responders and slow responders, through a median split of their FOK reaction times. We looked at whether fast responders performed differently compared to slow responders. Results of the 2 (Word Type: Strong vs. Weak) X 2 (Timing: Restricted vs. Self-paced) Between Subjects ANOVA on mean FOK ratings did not revealed any group differences (all *p*s >.53). We also compared FOK ratings for the fastest (the 25th percentile or lower) and the slowest (75th percentile or higher) FOK trials, again based on the FOK response time. Similarly, results of the 2 (Word Type: Strong vs. Weak) X 2 (Timing: Restricted vs. Self-paced) Between Subjects ANOVA did not revealed any significant findings (all *p*s > .79). Additionally, we furthered our investigation in a trial basis. We detected 7 participants who responded faster (<2000ms) in more than half of the trials. We compared their mean and median FOK magnitude ratings across slowest and fastest trials; yet, we could not find any significant pattern.

4.2.2.2 FOK accuracy

As in Experiment 2, results of 2 (Word Type: Strong vs. Weak) X 2 (Timing: Restricted vs. Self-paced) Between Subjects ANOVA on FOK gamma calculations revealed no effect of FOK response timing, word type and no interaction ps > .08 (see Table 4).

Table 4. Descriptives for Memory and Metamemory Measures

	Restricted		Self-Paced		Main Effects									
	Weak	Strong	Weak	Strong	Timing			Word Type			Interaction			*MSE*
	N = 13	*N* = 15	*N* = 13	*N* = 10										
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*	*F*	*p*	η^2_p	*F - t*	*p*	η^2_p	*F*	*p*	η^2_p	
Reaction Time (ms)	1653 (24)	1532 (47)	3705 (342)	3018 (191)										
Memory measures														
Recall (%)	0.37 (0.2)	0.33 (0.14)	0.27 (0.14)	0.26 (0.19)	-	-	-	0.387	0.704	-	-	-	-	-
Recognition (%)	0.8 (0.1)	0.79 (0.14)	0.74 (0.16)	0.73 (0.22)	1.803	0.187	0.044	0.043	0.837	0.001	0.015	0.904	0	0.025
Conditional recognition (%)	0.45 (0.15)	0.47 (0.1)	0.48 (0.12)	0.47 (0.15)	0.139	0.711	0.004	0.078	0.781	0.002	0.147	0.793	0.004	0.017
Metamemory measures														
FOK Mean	62.8 (11.05)	52.16 (17.77)	58.94 (13.95)	53.26 (24.27)	0.066	0.799	0.002	2.307	0.137	0.056	0.214	0.646	0.005	306
Correctly Recalled	88.83 (10.68)	80.49 (18.16)	98.73 (2.22)	94.09 (9.29)	9.841	0.003*	0.201	3.006	0.091	0.072	0.243	0.625	0.006	148
Incorrectly Recalled	38.67 (13.02)	31.95 (20.03)	44.38 (13.46)	37.87 (28.2)	0.925	0.342	0.023	1.197	0.281	0.03	0	0.986	0	388
Correctly Recognized	66.6 (11.64)	55.78 (18.18)	67.57 (12.65)	58.24 (22.4)	0.109	0.743	0.003	3.773	0.059	0.088	0.021	0.887	0.001	285
Incorrectly Recognized	25.03 (14.48)	18.81 (18.66)	33.19 (19.2)	34.75 (30.13)	0.128	0.722	0.003	3.423	0.072	0.081	0.356	0.554	0.009	450
FOK Gamma	0.45 (0.33)	0.47 (0.47)	0.7 (0.18)	0.61 (0.34)	3.146	0.084	0.075	0.095	0.759	0.002	0.291	0.593	0.007	0.125
FOK Gamma Inc Recog	0.17 (0.53)	0.29 (0.58)	0.55 (0.25)	0.3 (0.65)	1.468	0.233	0.036	0.157	0.694	0.004	1.34	0.254	0.033	0.28

Notes.
Significant findings are indicated with *
The t score within the word type main effects is only for the recall accuracy. We did not compare the effect of timing for recall accuracy because timing manipulation was not present during recall.

4.2.3 Partial information

We coded four different measures for partial information answers of the participants. First of all, we calculated number of partial information given for each target word. Second, we coded whether this information was phonological or semantic. Phonological information was related to qualities such as the first letter of the target, number of syllables in the target word, etc. Semantic information, on the other hand, was related to the meaning of the target, such as a word associated with the target, relationship between the target and the cue word, etc. Lastly, we coded whether participants mentioned the first association of the target word as a piece of partial information. We compared these four measures across word types. Participants reported first associations as partial information for the strong words (M = .65; *SD* = 1.15) significantly more than for the weak words (*M* = .05; *SD* = .22), *t* (41) = 2.296, *p* = .027, *r* = 34. However, we found no effect of partial information count, count of phonological partial information and semantic partial information across the two lists, *p*s > .3. See Table 5.

Table 5. Descriptives for the Partial Information Coding

	Weak		Strong			
	N = 20		*N* = 23			
	M	*SD*	*M*	*SD*	*t*	*p*
Total Partial Information	18.10	8.45	21.96	16.78	0.929	0.358
Phonological Partial Information	3.70	5.74	3.39	3.97	0.207	0.837
Semantic Partial Information	14.4	8.29	18.57	17.33	0.980	0.333
First Association Word	0.05	0.22	0.65	1.15	2.296	0.027

When we looked at the correlations between FOK magnitude ratings of self paced conditions and count of partial information we found that FOK magnitude ratings positively correlated with number of overall partial information (*r* = .552, *p* = .012) and semantic partial information (r = .608, p = .004). However in the restricted time conditions, there was a negative correlation between FOK magnitude and overall number of partial information (r = -.443, p = .034), semantic partial information (r = -.513, p = .012) and a positive correlation between phonological partial information (r = -.455, p = .029).

4.3 Discussion

The purpose of the Experiment 3 is to collect data on partial information of the target words provided by participants themselves. Results of the partial information coding revealed that participants explicitly discriminated between the target words. Participants gave more first association words as partial information for strong words. Moreover, although results are not significant, participants tend to give higher number of total partial information and semantic partial information for the strong words. Thus, this finding strengthens our argument that strong words do have higher levels of partial information compared to weak words. Moreover, we found that overall partial information count is positively correlated with FOK judgment magnitude for self paced trials, however, negatively correlated for restricted time conditions. Phonological partial information was found to be positively correlated with FOK judgments for restricted time conditions. These results highlight the fact that people might access different levels of partial information in different timing conditions.

As in previous experiments we investigated FOK accuracy, FOK magnitude for all judgments as well as FOK magnitude of both correctly recalled and incorrectly recalled trials. Rationale for the latter investigation is due to the evidence of the possible strategy use of recall phase, which always precedes the FOK judgment phase (see

Schwartz, Boduroglu & Tekcan, 2015). Results of FOK magnitude were as follows: When people correctly recalled the words they gave higher FOK responses in self-paced conditions, this finding was seen for both strong and weak words. Although there was an interaction for the overall mean FOK ratings and mean FOK ratings of incorrectly recalled judgments, this effect was not significant for Experiment 3. This is one of the major differences between Experiment 2 and Experiment 3.

The reason we were not able to replicate the pattern observed in Experiment 2 in this experiment may be partly due to the limited power given the small number of participants in Experiment 3. We compared the mean FOK ratings of each condition in Experiment 3 ($N = 57$) with the confidence interval scores of the mean FOK ratings in Experiment 2 ($N = 106$). None of the ratings, such as recall accuracy, recognition accuracy, FOK magnitude, FOK gamma, exceeded confidence intervals. Thus if we had a similar number of participants we may have replicated the results of Experiment 2. It is also possible that methodological differences between the Experiment 2 and Experiment 3 contributed to the results. Specifically, in Experiment 3 participants were instructed to give partial information regarding the target words just before the recognition phase. Thinking about a word and elaborating on it could be beneficial for future recognition. It is known that spaced learning is one of the most robust effects in cognitive psychology (e.g. Cain & Willey, 1939) yet it might also have drawbacks. Rethinking of a word could foster learning of that particular word already if participant remembers it. This may in turn increase recognition performance however, other possibilities could exist. For example, retrieval of a certain word could create chaining of associations which in turn may cause one to remember unrelated partial information (Nelson, McEvoy, Dennis, 2000). This may in turn mislead recognition performance. In this experiment this is particularly an issue because one of the lures in the recognition task was chosen to be the first association of the target from the available association norms. Data supports this interpretation in which, people make more recognition errors by choosing the first association in Experiment 3 (M = .77; SD = .16) compared to Experiment 2 (M = .06; SD = .06).

Moreover, according to part-set cuing, retrieval of any partial information related to the actual target word may have negative effects on future memory performance (Slamecka, 1968). Likewise, remembering a word that bears semantic relationship with actual target word may create retrieval induced forgetting and thus has negative impact on future memory performance (Anderson, Bjork, Bjork, 1994). These all would lower recognition accuracy in Experiment 3, which in turn could possibly reduce FOK accuracy. However, this was not the case; overall recognition accuracy in Experiment 2 was .80 and in Experiment 3 it was .76.

All in all, it is highly possible that recognition strategies in Experiment 3 were different than recognition strategies in Experiment 2. Nevertheless, we expect this methodological difference to effect FOK accuracy due to the possible effect on recognition phase not on FOK magnitude, recall and recognition accuracies which are preceded by PI phase.

CHAPTER 5

CONCLUSION

In the present thesis, the primary goal was to investigate whether qualitative aspects of words impact FOK judgments especially when time to make a judgment is either restricted or left open. More specifically, we investigated whether restricting the timing for the FOK judgments resulted in different FOK responses regarding target words with either strong or weak levels of association. We argued that levels of association could be used as a proxy for the accessibility of partial information, a concept frequently mentioned in the FOK literature (e.g. Koriat 1993; 1994; 1995) but without a clear operationalization. Effects of timing and levels of partial information had been tested by many studies in the literature; however, this is the first study to test this together and to come up with a comprehensive explanation. In this thesis, in particular in Experiment 2, we showed differential effects of FOK timing and levels of partial information of the target words on memory and metamemory measures.

Overall, in this study we replicated the effect of timing manipulation that was present in the literature (Benjamin, 2005; Koriat & Levy-Sadot, 2001; Thomas et al., 2011) in an episodic memory test. Under the self-paced condition, people rely on the accessibility of partial information related to the target words while making their FOK judgments. In this study we found that this effect is qualified by two different variables. One is the impact of recall accuracy: participants gave higher FOK judgments for the correctly recalled words in the self-paced timing condition compared to the restricted time condition. The second one is the level of partial information of the target words. In line with the literature, the more the partial information, the higher the FOK magnitude (Koriat, 1993; Schwartz et al., 2014). The novelty of our findings was to demonstrate this effect under different timing conditions. People give the highest FOK judgments for the strong words in self-paced timing condition. In this condition, people had the highest level of information (Koriat, 1993) and more time to evaluate the upcoming information (Thomas et al., 2011). We believe that this research is informative on the combined role of levels of partial information and timing of the FOK judgment on people's decision.

Particularly, in Experiment 1, we found that participants gave higher FOK judgments and greater recall accuracy for weak words compared to the strong words. These results were not in the expected direction because it is known from the literature that increases in the number of partial information increases the FOK judgment magnitude (see Koriat, 1993). In a follow up study we found that this effect is coming from the higher imagery ratings of the cue-target pairs in the weak condition. Given the importance of imagery in FOK judgments (Hertzog et al., 2014; Thomas et al., 2012) we wanted to eliminate this particular confound.

In Experiment 2, we rearranged the cue target pairs so that we eliminated the higher imagery between cue target pairs both in the weak and strong word groups. Results of the Experiment 2 revealed that when people correctly recalled the target words there was no word type effect, the only effect present was the effect of timing: participants gave higher ratings when they had longer time in the self-paced condition compared to the restricted time condition. Interestingly, when people incorrectly recalled the target words their FOK ratings did not differ across the two lists in the restricted time condition. However, in the self-paced condition their FOK ratings for strong words (i.e. levels of partial information) were higher compared to those for the weak words (i.e.

levels of partial information). This finding underlies the importance of two things. First, people use recall as an important cue for their FOK judgments and gave higher FOK ratings regardless of word type (see Schwartz et al., 2015). Second, participants used different strategies under restricted time conditions (Benjamin, 2005; Koriat & Levy-Sadot, 2001; Thomas et al., 2011) and the particular strategy relied upon depends on the qualitative aspects of the word list, in this particular case, the levels of partial information of the target (Thomas et al., 2012).

In Experiment 3, we investigated the explicit partial information responses of the participants. We asked participants to report as much information as possible regarding the target words. Results of the analyses revealed that participants significantly gave first association word of the targets as partial information in the strong word group. Moreover, we did not find evidence for the phonological, semantic and overall partial information for the strong word group compared to weak word group. However, when we compared the reported partial information with FOK magnitude ratings we found that there is a positive correlation for the semantic and overall partial information with FOK magnitude ratings in self paced condition and negative correlation for restricted time condition. In addition, phonological partial information positively correlated with FOK magnitude in restricted time conditions. These results highlight the fact that people might access different levels of partial information in different timing conditions.

It is highly possible that, for strong words, remembering a target word automatically activated the first association word of that target, which in turn activates a certain concept in memory. Thus, in such cases remembering the target word could act as remembering a cue, constructing fist associate as a target. In particular, remembering the target word acts as a cue to remember the fist association. Thus when people remember the first association (as if target) and remember back the actual target word (as if cue) automatically and make higher FOK judgments due to higher cue familiarity.

Lastly, we would like to mention about the possible drawbacks of this study. First of all, due to the restricted number of words in the Turkish Word Norms (Tekcan & Göz, 2006) we could not reach a desired number of target words fitting to strong and weak partial information criteria. Additionally, because words have intrinsic qualities such as imagery, concreteness and frequency it was possible that word lists contained confounds that we did not anticipate. We had to equalize the word groups in terms of above qualities, which resulted in a certain drop in the number of words. Future studies should test with higher number of words in each condition. Finally, we acknowledge the fact that this equalization process could have resulted in the loss of defining qualities of the words groups that we have created.

Future research is necessary to further investigate how partial information affects the nature of FOK judgments by manipulating using a different set of materials. This could be simply replicating the exact study without equalizing the words in terms of imagery and concreteness dimensions or using cue-target word pairs that qualify for different levels of partial information.

APPENDIX A

LISTS OF WEAK AND STRONG PARTIAL INFORMATION TARGET WORDS AND DESCRIPTIVES

Weak	Frequency	Imagery	Concreteness	SetSize	first Associate	
AYKIRI (AGAINST)	50	2.74	2.15	46	*TERS* (OPPOSITE)	10
BÖCEK (BUG)	37	6.32	6.75	44	*İĞRENÇ* (GROSS)	12
DELİK (HOLE)	111	5.65	6.09	50	*DÜĞME* (BUTTON)	11
ENGEL (OBSTUCTION)	126	3.97	4.42	43	*KOŞU* (RUN)	11
HIRÇINLIK (ACRIMONY)	6	3.21	2.10	41	*ASABİYET* (ANGER)	11
HORMON (HORMON)	59	3.02	5.36	48	*DOMATES* (TOMATO)	11
İP (ROPE)	96	6.25	6.84	47	*ÇAMAŞIR* (LAUNDRY)	11
KABİN (CABIN)	34	5.59	6.70	44	*SOYUNMAK* (UNDRESS)	10
KANCA (HOOK)	4	5.39	6.72	43	*BALIK* (FISH)	11
KARIŞIK (MIXED)	58	3.50	2.90	48	*SALATA* (SALAD)	10
KASNAK (TABORET)	2	4.02	6.36	44	*DİKİŞ* (STITCH)	12
MÜHENDİS (ENGINEER)	108	4.63	5.49	46	*CETVEL* (RULER)	11
OLAYLI (EVENTFUL)	1	3.24	2.61	47	*KAVGA* (FIGHT)	10
ÖLÜ (DEAD)	110	4.97	6.17	47	*CENAZE* (FUNERAL)	10
ÖLÜM (DEATH)	392	3.56	2.34	52	*MEZAR* (GRAVE)	9
ÖNEM (IMPORTANCE)	351	2.94	1.95	48	*DEĞER* (VALUE)	10
PATRON (BOSS)	102	5.36	5.81	48	*İŞ* (JOB)	12
ŞOK (SHOCK)	67	3.23	2.51	42	*ELEKTRİK* (ELECTICITY)	10
STANDART (STANDART)	153	2.86	2.47	45	*SAPMA* (DEVIATION9	10
SÜLÜN (PHEASANT)	4	2.45	5.53	45	*HAYVAN* (ANIMAL)	10
TAHTA (BOARD)	155	6.24	6.94	42	*MASA* (TABLE)	12
TIRTIK (SERRATION?)	1	3.41	5.32	45	*PÜRÜZ* (ROUGHNESS)	10
UÇAK (AEROPLANE)	194.00	6.47	6.91	39	*GÖKYÜZÜ* (SKY)	12
UYUMLU (COMPATIBLE)	55	2.89	2.10	50	*AHENK* (HARMONY)	7

Strong	Frequency	Imagery	Concreteness	SetSize	first Associate	
BULUT (CLOUD)	66	6.45	6.29	14	*YAĞMUR* (RAIN)	55
BRÜT (GROSS)	6.00	2.39	2.99	16	*NET* (TAKE HOME-SALARY)	54
ÇERÇEVE (FRAME)	179	6.11	6.76	14	*RESİM* (PICTURE)	60
DAMLA (DROP)	84.00	5.87	6.64	15	*YAĞMUR (RAIN)*	37
GÜNDÜZ (DAY)	57.00	4.94	5.27	15	*GECE* (NIGHT)	37
GÜNEY (SOUTH)	138	4.25	3.66	13	*KUZEY* (NORTH)	38
KADEH (GOBLET)	65	6.35	6.75	10	*ŞARAP* (WINE)	50
KAFİYE (RHYME)	14	4.27	3.93	11	*ŞİİR* (POEM)	42

KIŞ (WINTER)	123	5.83	5.15	13	*KAR* (SNOW)	48
KUYUMCU (JEWELLER)	15	6.02	6.36	11	*ALTIN* (GOLD)	84
LÂMBA (LIGHTBULB)	83	6.36	6.84	12	*IŞIK* (LIGHT)	74
LEHÇE (DIALECT)	1.00	3.27	3.91	15	*DİL* (LANGUAGE)	62
MISRA (VERSE)	20	4.75	5.31	9	*ŞİİR (POEM)*	78
MUSON (MUNSOON)	1	3.08	4.70	11	*YAĞMUR (RAIN)*	66
NEZLE (COLD)	13	5.18	5.20	13	*GRİP* (FLU)	42
ÖĞLEN (MIDDAY)	18.00	4.23	4.23	15	*YEMEK* (FOOD)	49
ÖNCÜLÜK (LEADERSHIP)	14	3.20	2.23	14	*LİDERLİK* (LIDERSHIP)	79
ÖTÜCÜ (SINGING)	1	2.62	3.96	13	*KUŞ* (BIRD)	79
PIHTI (CLOT)	8	4.40	6.33	8	*KAN* (BLOOD)	91
PİRİNÇ (RICE)	101.00	6.50	6.69	15	*PİLAV* (RICE)	74
RAKAM (NUMBER)	148	5.89	4.14	13	*SAYI* (NUMBER)	65
SEMER (PACK SADDLE)	1	4.38	6.42	12	*EŞEK* (DONKEY)	49
TÜTÜN (TOBACCO)	28	6.07	6.82	13	*SİGARA* (CIGARETTE)	84
ÜLSER (ULCER)	16	3.44	5.10	14	*MİDE* (STOMACH)	58

APPENDIX B

LIST OF CUE WORDS AND DESCRIPTIVES

Cue Words	Frequency	Imagery	Concreteness	Set Size	first Association
METAL (METAL)	87	5.32	6.51	30	18
PALTO (COAT)	25	6.34	6.77	26	16
PROTEİN (PROTEIN)	64	3.29	5.17	25	13
BARIŞ (PEACE)	67	3.43	2.44	28	10
ELEŞTİRMEN (COMMENTATOR)	22	3.8	5.34	27	11
VALİ (GOVERNOR)	80	5.09	6.11	28	12
BELDE (TOWN)	24	3.71	5.35	28	14
İSHAL (DIARRHEA)	32	4.67	5.72	29	15
MAKYAJ (MAKEUP)	79	5.39	5.93	27	14
ÖZLEM (LONGING)	73	3.17	1.64	29	16
TELÂŞ (HURRY)	58	4.16	2.49	26	13
ÇİRKİN (UGLY)	63	4.08	2.61	27	10
KİMYA (CHEMISTRY)	50	3.58	4.49	26	15
SEÇMEN (VOTER)	52	4.82	6.2	26	8
EŞSİZ (UNIQUE)	36	2.16	2.35	29	17
TREN (TRAIN)	83	6.45	6.9	25	11
NESİL (GENERATION)	64	2.71	3.32	29	16
SOMUT (CONCRETE)	64	3.35	3.18	30	15
TAVUK (CHICKEN)	92	6.69	6.88	27	13
MİNERAL (MINERAL)	52	3.23	5.47	26	12
SALATA (SALAD)	72	6.51	6.79	25	16
DRAM (DRAMA)	50	2.82	2.84	25	11
MUZ (BANANA)	67	6.73	6.92	25	9
PASTA (CAKE)	79	6.48	6.77	28	12

APPENDIX C

RECOGNITION LURES FOR THE WEAK AND STRONG WORD SETS

Recognition Lure	Frequency	Imagery	Concreteness	Set Size	first Associate
BORÇLU (OWING)	32.00	3.08	4.32	30	18
CESUR (BRAVE)	39.00	3.49	2.25	30	9
ÇIĞLIK (CRY)	74.00	4.60	5.20	30	11
DUA (PRAY)	51.00	4.00	2.84	30	15
EŞEK (DONKEY)	52.00	6.40	6.80	30	13
FORMÜL (FORMULA)	50.00	4.08	4.58	29	16
GİYİM (CLOTHING)	74.00	5.71	5.45	30	14
GÖMLEK (SHIRT)	73.00	6.64	6.89	30	13
KARTON (CARTON)	22.00	5.46	6.77	30	13
KORKUNÇ (SCARY)	91.00	3.48	2.30	30	17
LÜKS (LUXURIOUS)	91.00	3.88	2.89	30	18
MİRAS (LEGACY)	52.00	3.64	5.25	29	11
MORAL (MORALE)	55.00	3.08	1.84	28	23
ÖZET (SUMMARY)	68.00	3.92	4.73	30	11
PİLÂV (RICE)	53.00	6.58	6.85	28	12
REZALET (DISGRACE)	31.00	2.60	2.69	30	19
SAHİL (BEACH)	73.00	6.29	6.38	29	12
SÜRPRİZ (SURPRISE)	49.00	3.43	2.71	28	11
UĞULTU (HUMMING)	23.00	3.56	4.77	30	12
UZAY (SPACE)	62.00	4.89	4.64	29	16
VATAN (HOMELAND)	58.00	4.06	4.43	28	10
YAKIT (FUEL)	72.00	4.98	6.48	28	16
YORGUNLUK (TIREDNESS)	63.00	4.20	2.80	28	12
ZİHİN (MIND)	77.00	2.91	2.50	29	11

Recognition Lure	Frequency	Imagery	Concreteness	Set Size	first Associate
ATIK (WASTE)	49.00	3.81	6.04	27	15
ATOM (ATOM)	52.00	3.16	4.73	28	14
BALKON (BALCONY)	57.00	6.45	6.87	28	18
BOYA (PAINT)	98.00	5.35	6.43	26	14
ÇUKUR (HOLE)	70.00	5.93	6.39	26	18
ESİR (CAPTIVE)	30.00	4.64	5.06	27	16
EVLÂT (CHILD)	78.00	5.29	6.00	25	11
GARSON (WAITER)	90.00	6.24	6.61	26	15
HARİTA (MAP)	61.00	6.26	6.70	26	10
HASAR (DAMAGE)	62.00	3.94	4.47	27	10
İSKELET (SKELETON)	26.00	6.03	6.72	27	13
KAHVALTI (BREAKFAST)	71.00	6.25	6.37	25	8
KÖPÜK (BUBBLE)	54.00	5.78	6.49	25	13
MİNDER (CUSHION)	24.00	6.31	6.86	26	13
MOBİLYA (FURNITURE)	89.00	6.14	6.76	25	11
OT (GRASS)	70.00	6.29	6.79	25	8
PRENSİP (PRINCIPLE)	61.00	2.85	1.98	26	11
SEYİRCİ (SPECTATOR)	85.00	6.16	6.47	27	17
SEZON (SEASON)	66.00	2.76	3.59	26	12
TAKSİT (INSTALLMENT)	89.00	3.92	4.58	27	14
TELÂŞ (HURRY)	58.00	4.16	2.49	26	13
TUTKU (PASSION)	62.00	3.85	2.14	27	9
UYGARLIK (CIVILIZATION)	61.00	3.81	3.10	27	8
ZEKÂ (INTELLIGENCE)	66.00	2.92	2.33	26	13

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.

Benjamin, A. S. (2005). Response speeding mediates the contribution of cue familiarity and target retrievability to metamnemonic judgments. *Psychonomic Bulletin & Review*, *12*, 874-879.

Bjork, R. A., & Benjamin, A. S. (2011). On the symbiosis of remembering, forgetting, and learning. *Successful remembering and successful forgetting: A festschrift in honor of Robert A. Bjork*, 1-22.

Bjork, R. A., & Bjork, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. *From learning processes to cognitive processes: Essays in honor of William K. Estes*, *2*, 35-67.

Boduroglu, A., Pehlivanoglu, D., Tekcan, A. İ., & Kapucu, A. (2015). Effects of self-referencing on FOK accuracy and recollective experience. *Memory*, *23*, 736 – 747.

Brewer, G. A., Marsh, R. L., Clark-Foos, A., & Meeks, J. T. (2010). Noncriterial recollection influences metacognitive monitoring and control processes. *The Quarterly Journal of Experimental Psychology*, *63*(10), 1936-1942.

Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological bulletin*, *109*(2), 204.

Cain, L. F., & Willey, R. D. V. (1939). The effect of spaced learning on the curve of retention. *Journal of Experimental Psychology*, *25*(2), 209.

Dunlosky, J., & Metcalfe, J. (2008). Metacognition. Sage Publications.

Galbraith, R. C., & Underwood, B. J. (1973). Perceived frequency of concrete and abstract words. *Memory & Cognition*, *1*(1), 56-60.

Gonzalez, R., & Nelson, T. O. (1996). Measuring ordinal association in situations that contain tied scores. *Psychological bulletin*, *119*(1), 159.

Göz, İ. (2003). *Yazılı türkçenin kelime sıklığı sözlüğü* (Vol. 823). Türk Dil Kurumu.

Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology*, *56*(4), 208.

Hertzog, C., Dunlosky, J., & Sinclair, S. M. (2010). Episodic feeling-of-knowing resolution derives from the quality of original encoding. *Memory & Cognition*, *38*, 771-784.

Hertzog, C., Fulton, E. K., Sinclair, S. M., & Dunlosky, J. (2014). Recalled aspects of original encoding strategies influence episodic feelings of knowing. *Memory & cognition, 42*(1), 126-140.

Hosey, L. A., Peynircioğlu, Z. F., & Rabinovitz, B. E. (2009). Feeling of knowing for names in response to faces. *Acta Psychologica*, *130*(3), 214-224.

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review*, *100*(4), 609.

Koriat, A. (1994). Memory's knowledge of its own knowledge: The accessibility account of the feeling of knowing. In J. Metcalfe & A. Shimamura (Eds), *Metacognition: Knowing about knowing* (pp. 115-135). Cambridge, MA, US: The MIT Press.

Koriat, A. (1995). Dissociating knowing and the feeling of knowing: Further evidence for the accessibility model. *Journal of Experimental Psychology: General*, *124*(3), 311.

Koriat, A., & Levy-Sadot, R. (2001). The combined contributions of the cue-familiarity and accessibility heuristics to feelings of knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(1), 34.

Koriat, A., Levy-Sadot, R., Edry, E., & de Marcas, S. (2003). What do we know about what we cannot remember? Accessing the semantic attributes of words that cannot be recalled. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(6), 1095.

Lupker, S. J., Harbluk, J. L., & Patrick, A. S. (1991). Memory for things forgotten. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 897-907.

Marschark, M., & Hunt, R. R. (1989). A reexamination of the role of imagery in learning and memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(4), 710.

Metcalfe, J., Schwartz, B. L., & Joaquim, S. G. (1993). The cue-familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(4), 851.

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological bulletin*, *95*(1), 109.

Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist*, *51*(2), 102.

Nelson, D. L., McEvoy, C. L., & Dennis, S. (2000). What is free association and what does it measure?. Memory & cognition, 28(6), 887-899.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 402-407.

Nelson, D. L., Schreiber, T. A., & McEvoy, C. L. (1992). Processing implicit and explicit representations. *Psychological review*, *99*(2), 322.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and some new findings. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation* (pp. 125-173). New York: Academic Press.

Nelson, T. O., Gerler, D., & Narens, L. (1984). Accuracy of feeling-of-knowing judgments for predicting perceptual identification and relearning. *Journal of Experimental Psychology: General*, *113*(2), 282.

Pansky, A., & Koriat, A. (2004). The basic-level convergence effect in memory distortions. *Psychological Science*, *15*(1), 52-59.

Psychology Software Tools, Inc. [E-Prime 2.0]. (2012). Retrieved from http://www.pstnet.com.

Reder, L. M. (1987). Strategy selection in question answering. *Cognitive Psychology*, *19*(1), 90-138.

Sacher, M., Taconnat, L.,& Souchay, C. (2009).Divided attention at encoding: Effect on feeling-of-knowing. *Consciousness & Cognition*, *18*, 754-761.

Schacter, D. L. (1983). Feeling of knowing in episodic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*, 39-54.

Schwartz, B. L. (1993). Mechanisms of feeling of knowing in memory (Unpublished doctoral dissertation).

Schwartz, B. L., Boduroglu, A., & Tekcan, A. İ. (2015). Methodological concerns: the feeling-of-knowing task affects resolution. *Metacognition and Learning*, 1-12.

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*, 1074-1083.

Schwartz, B. L., & Smith, S. M. (1997). The retrieval of related information influences tip-of-the-tongue states. *Journal of Memory and Language*, *36*(1), 68-86.

Schwartz, B. L., Pillot, M., & Bacon, E. (2014). Contextual information influences the feeling of knowing in episodic memory. *Consciousness and Cognition*, *29*, 96-104.

Slamecka, N. J. (1968). An examination of trace storage in free recall. *Journal of experimental psychology*, *76*(4p1), 504.

Tekcan, A. I., & Akturk, M. (2001). Are you sure you forgot? Feeling of knowing indirected forgetting. *Journal of Experimental Psychology: Learning, Memory,and Cognition*, *27*, 1487-1490

Tekcan, A. I., & Göz, I. (2005). *Turkce Kelime Normlari (Turkish Word Norms).* Istanbul: Bogazici Universitesi Yayinevi.

Thomas, A. K., Bulevich, J. B., & Dubois, S. J. (2011). Context affects feeling-of-knowing accuracy in younger and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(1), 96.

Thomas, A. K., Bulevich, J. B., & Dubois, S. J. (2012). An analysis of the determinants of the feeling of knowing. *Consciousness and Cognition*, *21*(4), 1681-1694.