Cognitive control of the contents of working memory

By

EDA MIZRAK

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Psychology

in the Graduate School of Social Sciences and Humanities

of

Koç University

January 2017

Koc University

Graduate School of Social Sciences and Humanities

This is to certify that I have examined this copy of a Ph.D. dissertation by

Eda Mızrak

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Committee Members:

İlke Öztekin, Associate Professor, Koç University

Fuat Balcı, Associate Professor, Koç University

Sami Gülgöz, Professor, Koç University

Ümit Akırmak, Assistant Professor, Istanbul Bilgi University

Ayşecan Boduroğlu, Associate Professor, Boğaziçi University

Date: 11/01/2017

THESIS ABSTRACT

Working memory (WM) is a dynamic system that allows us to actively maintain recently encountered information so that it is available for ongoing cognitive operations. Information held in an accessible state in WM is used to guide many complex decisions and behavior. However, WM has capacity limitations. Efficient use of this capacity-limited system is the key to successful execution of complex tasks and requires control over the contents maintained in WM. There are several processes that control WM such as updating active content according to task goals, focusing attention on the relevant content, removing irrelevant information, and controlled retrieval of relevant information. In this dissertation, I focus on one of these processes: controlled retrieval of relevant information.

The set of studies conducted in this dissertation use convergent methodologies to test possible factors that influence successful controlled retrieval. Chapters I and II introduce two behavioral studies employing the speed-accuracy-trade-off (SAT) procedure to independently measure the speed and accuracy of working memory retrieval. Chapter III introduces a study that employs functional magnetic resonance imaging (fMRI) to further investigate the findings from Chapter II and complement the behavioral measures with neural measures.

Chapter I examines how individual differences in working memory capacity (WMC) is an important indicator of the efficiency of controlled retrieval from working memory. Individuals who fell in the upper and lower quartile on a WMC measurement task were to retrieve temporal order information of the recently presented items, which required a controlled serial search through the contents of WM. The prediction was that those who rank higher on WMC measures would perform well at this search, which would result in both faster and more accurate retrieval when compared to those with lower WMC. Findings from this chapter revealed that WMC was indeed successful at predicting retrieval efficiency such that individuals who fell in upper and lower quartile in the WMC measurement task differed in both speed and accuracy while retrieving temporal order information from WM.

In Chapters II and III, I considered another possible factor that might affect retrieval efficiency: affective content of to be remembered material. I used two different tasks in which controlled retrieval demands were increased by manipulating the presence of proactive interference (PI). Participants needed to retrieve additional diagnostic information about the test item either at the item level or context level to correctly recognize it in the presence of interference. I evaluated the differences between neutral and emotional materials, foreseeing possible differences in the process in which PI builds up and the process in which it is resolved via controlled retrieval. There was less PI build up for emotional materials in both tasks, but controlled retrieval was less efficient if resolving PI required context retrieval; on the other hand, if participants could recover relevant information by retrieving details about the item, there was no such differences in controlled retrieval across emotional and neutral information.

Together, these Chapters demonstrate that the efficiency of controlled retrieval depends on (1) an individual’s ability to maintain relevant materials in an accessible state while concurrently executing a task (i.e., greater WMC) and (2) whether the encountered information is emotional or neutral. Studies presented here add up to a growing literature of studies investigating how we can utilize use of working memory which determines our intellectual abilities.

TEZ ÖZETİ

Çalışma belleği (ÇB) dinamik bir bellek sistemi olmakla beraber yakın zamanda öğrenilmiş bilgilerin aktif ve kullanıma hazır bir şekilde tutulmasını sağlar. ÇB'de erişilebilir bir durumda tutulan bilgiler bir çok karar verme ve davranış mekanizmalarını yönlendirmek için kullanılmaktadır. Ancak ÇB kısıtlı bir kapasiteye sahiptir. Kısıtlı kapasiteye sahip bu bellek sisteminin verimli olarak kullanılabilmesi, ÇB'te tutulan bilgilerin kontrolünü gerektiren komplike davranış ve durumların başarılı bir şekilde idame ettirilebilmesinde büyük önem taşımaktadır. ÇB'nin içeriğinin kontrolüne katkıda bulunan birçok süreç vardır. Örneğin, aktif içeriğin amaçlara göre yenilenmesi, işe yaramayan bilgilerin çıkarılması, ve daha sonra kullanılacak içeriğin kontrollü olarak bellekten geri çağırılması gibi. Bu tez gerekli bilgilerin bellekten kontrollü olarak çağırılma süreci üzerine odaklanmaktadır.

Burada sunulan çalışmalar birbirini tamamlayıcı metodlar içermekte olup kontrollü geri çağırım sürecini etkileyen olası etkenleri sınamaktadır. Birinci ve ikinci bölüm tepki-sınırı hız-başarı ödünleşimi yöntemi içermekte olup hız ve doğruluk sürecini birbirinden bağımsız olarak incelemeyi hedeflemiştir. Bölüm III ise beyin görüntüleme yöntemi kullanarak Bölüm II'deki davranışsal bulguları tamamlamaya yönelik yapılmıştır.

Birinci bölüm çalışma belleği kapasitesindeki bireysel farklılıkların ÇB'den kontrollü geri çağırma sürecini nasıl etkilediğini araştırmaktadır. ÇB kapasitesini ölçen bir ölçekle test edilen bir örneklemin üst ve alt uzamına düşen bireyler çalışma belleğindeki bilgilerin zamansal olarak geri çağırılmasını gerektiren bir deneyle test edilmiştir. Zamansal bilginin geri çağırımı ÇB'deki içeriğin kontrollü bir şekilde aranmasını gerektirir. Bu çalışmanın bulgularıyla ilgili hipotezimiz yüksek uzam bireylerin bellekten geri çağırılacak bilgiyi arama sürecindeki performanslarının düşük uzamdaki bireylerden daha verimli olacağına yöneliktir. Buna ek olarak, performans farklılıklarının sadece doğruluk üzerine değil hız üzerine de yansıyacağı öngörülmüştür. Bu bölümde yaptığımız çalışmanın sonuçları hipotezlerimizle

örtüşmekte olup, düşük uzam bireylerin ÇB'de bulunan öğelerin zamansal bilgilerine erişmek için yapılan kontrollü arama sürecinde yüksek uzam bireylere kıyasla daha az verimli olduğu gözlemlenmiştir.

İkinci ve üçüncü bölüm, kontrollü geri çağırım sürecini farklı bir etkenle ölçmeye yönelik yapılmıştır. Geri çağırım sürecinde bozucu etki bulunması hafıza performansını kötü yönde etkilemekte olup, katılımcıların bozucu etkiyi çözümleyip öğeleri tanıyabilmeleri için bellekte bulunan doğru kaynak ya da öğe bilgisine kontrollü bir sekilde erişebilmeleri gerekmektedir. Bu tezde geri çağırım ihtiyaçlarını değiştirmeye yönelik iki farklı bozucu etkiyi kontrol eden yöntem kullanılmıştır. Bölüm II'de kullanılan yöntemde bozucu etkinin etkili bir şekilde çözünümü doğru kaynak (context) bilgisine erişime dayanmaktadır. Bölüm III'te ise bozucu etkinin çözünümü doğrudan öğe seviyesinde bilginin geri çağırımına dayanmaktadır. Bu iki çalışmada da nötr ve duygusal içerikli materyallerin hatırlanma performansı arasındaki farklılıklar incelenmiştir. Her iki çalışmada da bulgular şuna işaret etmektedir: bozucu etkinin birikim düzeyi duygusal öğeler için nötr öğelere kıyasla daha azdır. Bununla birlikte, duygusal içeriğin kontrollü süreçlere etkisi ÇB'den geri çağırılacak bilginin öğe ya da kaynak seviyesinde olmasına bağlı olarak değişmektedir. Geri çağırılacak bilgi kaynak bilgisi ise kontrollü süreçler duygu içeriğinden negatif yönde etkilenmiştir. Diğer yandan eğer geri çağırılacak bilgi öğe seviyesinde ise duygu içeriğinin pozitif etkisi olduğu görülmüştür.

Tezde yer alan bölümler bütün olarak bellekten kontrollü geri çekim sürecinin verimliliğinin (i) kişilerin daha sonra kullanılacak ilgili materyalleri başka bir görev yaparken erişilebilir bir seviyede tutma yetisi ve (ii) sunulan bilgilerin duygu içeriği tarafından etkilendiğini göstermektedir. Burada sunulan çalışmalar ÇB'nin hangi koşullarda verimli kullanıldığını araştıran bir literatüre katkıda bulunmaktadır.

ACKNOWLEDGEMENTS

First and foremost I want to thank my advisor Ilke Oztekin for giving me the opportunity to explore my research interests. Without her guidance and encouragement, I would not have accomplished what I have.

I also thank Charan Ranganath for giving me the opportunity to pursue my research at the University of California, Davis, and sharing his wisdom with me. This visit would not have been possible without the generous support of The Scientific and Technological Research Council of Turkey (TÜBİTAK), specifically the Doctoral Research Scholarship I received. The work performed in Davis contributed to the third chapter of this dissertation.

I thank my partner in crime, Henrik Singmann whom I met during a conference presentation of the study in Chapter I, for his endless support, encouragement, and patience. He always pushed me to do better. I am grateful for being with such an extraordinary person.

I would like to thank my house mates; Filiz, Hilal, Ipek, and my office mates Ceyla and Deniz for sharing the best and the worst of grad school with me. You gave me the courage to go on. I cannot thank you enough for listening to my complaints and putting up with me during this PhD. We had the worst and best together.

I also thank all the members of Dynamic Memory Lab and my dear friends Quentin, Fiona, Tamar, Akiva, Ena, Joel, Claudia, and Hugo whom I have met in Davis for making my Davis life fun and memorable. Nichole and Tanya, you guys are my rock. I do not know how I would have written this dissertation without your support.

I thank my amazing friends, Ayşen, Eleni, Kerem, and Merve for being in my life ever since I can remember. I thank you for “forcing” me to get out of my study dome and have fun. Knowing that I would meet you at the weekends made me go through the 24 hour work days. I never had an ordinary day with you and may years bring more phenomenal memories to our lives.

I am utterly grateful for my friends Dilek, Emel, Fisun, Merve, and Zeliha for always being there. Emel, knowing that you are there – no matter what – makes my life cool as a cucumber. Our mutual understanding is priceless.

Last, I thank my mother, father, and most importantly my sister for being there for me at all costs. Being in a family with such strong women characters and a father who always supported our empowerment made me the woman I am today. I am forever thankful for being a part of this family.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

THESIS INTRODUCTION

In this piece of work, we investigated the factors that affected the efficiency of controlled retrieval from working memory. Working memory (WM) is the cognitive system that mediates the short-term maintenance and manipulation of the information held in an accessible state to be used for an ongoing task. Put simply, it can be thought as a mental blackboard where we write down the information we will be using interactively during a cognitive task, such as keeping an address description in mind until we reach our destination. WM serves as a mental workspace that information relevant to a task is kept accessible, however, its capacity is limited. The ability to hold relevant information accessible determines our performance in many cognitive tasks and has been shown to correlate with general fluid intelligence (Conway, Kane, & Engle, 2003). WM capacity is a limit on the complexity of our thoughts and our cognitive abilities (Oberauer, Farrell, Jarrold, & Lewandowsky, 2016). Moreover, deficits in WM and depletion of WM capacity are mentioned to have a role in many psychological disorders (e.g. depression), social constructs (e.g. stereotype threat, prejudice) and neuropsychological disorders (e.g. Alzheimer's disease) (reviewed in Ilkowska & Engle, 2010). Therefore, understanding how we efficiently retrieve relevant information from WM is an essential step toward understanding how we are limited in our cognitive abilities and why we differ in these abilities.

Efficient use of a limited resource such as WM requires the cognitive control over the actively maintained contents in WM. Many complex tasks demand both the storage and processing of information concurrently, which is moderated by selectively attending to the elements on the mental blackboard. We bring the items that are needed for immediate processing to our focus of attention (FoA). Previous research has shown that the item which is currently in the FoA is in a readily accessible state while the rest of the information held available in the blackboard needs to be shunted to the focus of attention when needed (reviewed in McElree, 2006). Successfully shunting the required/relevant information to the FoA might depend on a) the ability to maintain relevant information in an accessible state b) Protecting the WM contents from the interference of irrelevant information c) Selective/controlled retrieval of the relevant information d) Switching our attention to the retrieved content.

This dissertation focuses on one of the cognitive control processes that is needed to utilize WM use; selective/controlled retrieval of relevant information from working memory. It consists of three studies that aimed to reveal how the efficiency of controlled retrieval is affected by a) working memory capacity (WMC), b) the type of information (neutral vs emotional) that needs to be retrieved. WMC is operationally defined as the number of items that can be recalled during a complex span task (Barrett, Tugade, & Engle, 2004) which measures an individual's ability to maintain representations in an accessible state to use later while executing a task. In the first chapter, we investigated the relationship between individual differences in WMC and efficiency of controlled serial search through contents of WM to retrieve temporal order information. The remaining two studies focused on how stimuli type (i.e., emotional vs. neutral) affects the controlled retrieval in the presence of interference from previously learned yet task-irrelevant information. Contents of working memory are constantly subject to interference from previous encounters, a phenomenon known as proactive interference (PI). Experimental psychologists showed that presence of PI leads to noticeable memory failures in both short-term and long-term memory (Jonides & Nee, 2006). It has been suggested that recovering from the detrimental effects of PI requires engaging with controlled/selective retrieval (Badre & Wagner, 2005; Nyhus & Badre, 2015). We entertained the possibility that emotional previous encounters might have a different interference effect on retrieval from WM compared to non-emotional encounters.

All three studies used two choice tasks, either two-forced alternative or binary, in which three to five items were presented in a to-be-remembered study list and participants were either asked to indicate to recognize a test probe or judge the recency of two test probes. Our preferred procedure in chapters I and II and choice of modeling approach in chapter III were employed to a) to have independent measures of retrieval speed and retrieval accuracy and use retrieval speed as an index of efficiency of controlled retrieval b) to fully capture the impact of our manipulations on latent processes that contribute to performance by accounting for slow and fast correct/incorrect responses.

It is a well investigated phenomenon that people trade accuracy for speed or vice versa (reviewed in Wickelgren, 1977). When people speed up while making decisions in experimental tasks, we usually observe more errors and when they slow down accuracy improves. However, sometimes, often in certain memory tasks (i.e. recent-probes task), the correct response can be given by either relying on fast/automatic assessments or by engaging with controlled operations. In some cases, a fast/automatic response can lead to both fast correct and incorrect responses. In other cases, when the correct response requires controlled operations, a correct response will take longer whereas if information required is not accessible even after engaging with controlled operations, this will result in slow errors. The tasks we used specifically demanded individuals to access information needed by selective/controlled retrieval of relevant information hence we used mathematical modeling to account for all of the response types.

Collectively, this dissertation reveals important factors that alters the efficiency of controlled/selective retrieval from working memory through the use of mathematical modeling and brain imaging methods. First, by investigating the full time course of retrieval and measuring retrieval speed and accuracy independently we were able to show how individual variations in WMC and type of information held in WM impacted controlled retrieval efficiency. Second, by our joint modeling approach which combined neural measures with the estimates from our model fittings we provided evidence for how retrieval efficiency changes as a function of proactive interference for different types of stimuli. Our studies aim to further the field of working memory research by investigating/testing unanswered questions regarding factors that affect controlled retrieval by employing elegant designs and advanced methodologies and analysis techniques.

CHAPTER I

Working Memory Capacity and Controlled Serial Memory Search

1.1.Abstract

The speed–accuracy trade-off (SAT) procedure was used to investigate the relationship between working memory capacity (WMC) and the dynamics of temporal order memory retrieval. High- and low-span participants (HSs, LSs) studied sequentially presented five-item lists, followed by two probes from the study list. Participants indicated the more recent probe. Overall, accuracy was higher for HSs compared to LSs. Crucially, in contrast to previous investigations that observed no impact of WMC on speed of access to item information in memory (e.g., Öztekin & McElree, 2010), recovery of temporal order memory was slower for LSs. While accessing an item's representation in memory can be direct, recovery of relational information such as temporal order information requires a more controlled serial memory search. Collectively, these data indicate that WMC effects are particularly prominent during high demands of cognitive control, such as serial search operations necessary to access temporal order information from memory.

1.2. Introduction

Individual variations in working memory capacity (WMC) correlate with performance in a broad range of complex cognitive activities such as reading comprehension (e.g., Daneman & Carpenter, 1980; Kane & McVay, 2012), logical reasoning (Kyllonen & Christal, 1990), drawing inferences (Linderholm, 2002) and retrieving relevant information from memory (Öztekin, & McElree, 2010). WMC is also found to be a good predictor of general fluid intelligence (Conway, Cowan, Bunting, Therriault, & Minkoff, 2002; Conway et al., 2003) and Scholastic Aptitude Test scores (Engle, Tuholski, Laughlin, & Conway, 1999). In addition to WMC differences predicting performance on cognitive function, WM deficits have been found to be related to psychological disorders such as schizophrenia, attention deficit disorder, and Alzheimer's disease as well (Ilkowska & Engle, 2010).

WMC can be measured by complex span (CS) tasks, and individual differences can then be examined by comparing performance of individual scoring in the upper and lower ends: high span individuals (HSs), who perform well in CS tasks and score in the upper quartile, and low span individuals (LSs), whose scores fall within the lower quartile (e.g. Kane & Engle, 2003; also see Redick et al., 2012 on the use of CS tasks to measure WMC). Numerous studies have compared the performance of HSs and LSs across multiple various tasks. Differences in performance have been noted even on tasks without an explicit memory component such as the dichotic listening task (Conway, Cowan, & Bunting, 2001), Stroop (Kane & Engle, 2003), the antisaccade (Kane et al., 2001), flanker (Redick & Engle, 2006), and go/no-go tasks (Redick, Calvo, Gay, & Engle, 2011). CS tasks operationally measure the number of items that can be recalled, however, it is thought to tap a domain-general construct that constitutes the strong correlation between CS performance and cognition (Broadway, Redick, & Engle, 2010).

It is crucial to note that these differences emerge under certain conditions when controlled attention is required to actively maintain task relevant information, especially in situations where there is substantial external and internal distraction. Theories that base attentional control as the underlying factor for WMC differences posit that the ability to maintain goal-relevant representations in the face of distraction requires successful and controlled allocation of attention (Engle, 2002, 2010; Engle & Kane, 2004). Accordingly, the controlled attention framework suggests that HSs are better at allocating their attention on goal-relevant information than LSs. Critically, this theory predicts that LSs perform worse in the presence of interference and distraction, but perform comparable in its absence, indicating that WMC does not reflect a general deficit in cognitive processing. In this respect, attentional control determines the predictive power of WMC on cognitive tasks.

More recently, studies considered WMC related effects using memory tasks that specifically measured retrieval differences across the high and low span groups. Unsworth and Engle (2007) proposed a “dual-component framework’’ for explaining individual differences in WMC, and suggested that differences between HSs and LSs arise from the maintenance of items in an accessible state and retrieval of items that are not in accessible state via controlled/strategic retrieval. Within this framework, while attentional control is still an important component, controlled retrieval is also an essential determinant of WMC (Unsworth & Spillers, 2010). Attentional control is required to maintain representations in an active state and successful allocation of attention to goal-relevant information protects the memory contents from interfering material. If external or internal distractors capture attention, the maintenance/availability of the representations would be affected. In this case, a controlled search through memory representations would be required to recover the items that were not maintained in an accessible state. Successful retrieval of items may then rely on the encoding quality, the ability to reinstate the context at retrieval, and delimit the search set to target items via excluding the interfering items (Unsworth & Engle, 2007). Accordingly, in this framework individual differences in controlled attention and controlled retrieval jointly explain the individual variations in WMC. While a majority of studies investigating individual differences in WMC focused on maintenance operations, the impact of WMC on controlled retrieval of task-relevant information is a relatively newly attended area of research that promises to improve our understanding regarding the relationship between WMC and cognition.

How does WMC affect the dynamics of memory retrieval? Öztekin and McElree (2010) tested HSs and LSs with a modified version of Sternberg probe recognition task. Because the traditional probe recognition task requires access to only item representations, this can be achieved via direct access to the relevant representation, without the necessity to engage in a search through memory representations (e.g., see McElree, 2006 for an overview). However, in the modified version (Monsell, 1978) on particular trials proactive interference was induced by selecting a lure from the previous study list. Due to the high residual familiarity of this lure, successful resolution of PI necessitates controlled processing, such as controlled episodic memory retrieval (e.g., Badre & Wagner, 2005; see Jonides & Nee, 2006 for review; Oztekin, Curtis, & McElree, 2009; Öztekin & McElree, 2007). Their findings showed although HSs exhibited higher accuracy than LSs, retrieval speed differences depended on whether the trials required controlled processing or not. Namely, LSs' speed of directly accessing the items from their memory was at similar levels with HSs in the absence of interference. However, when there was interference in the retrieval context, LSs were delayed in initiating the controlled retrieval operations that resolve interference in memory. Accordingly the results suggested that, with respect to access to item information in memory, WMC affected speed of processing only when the task demanded controlled retrieval operations due to the presence of interference in the retrieval context.

Manipulating interference in memory is one way to manipulate demands on controlled retrieval, and a well-established determinant of WMC related changes in cognitive performance (e.g., Kane & Engle, 2000; Öztekin & McElree, 2010). Another variable that determines the nature of retrieval operations is the type of information that needs to be accessed from memory. Specifically, access to item representations in memory can be achieved via a direct access mechanism, without the need to search through irrelevant memory representations (see McElree, 2006 for a review). Access to relational information (e.g., temporal or spatial order) on the other hand requires a slower, more controlled serial memory search, namely, controlled retrieval (Hacker, 1980; McElree & Dosher, 1993; Öztekin, McElree, Staresina, & Davachi, 2008). The present study aimed to assess the impact of WMC on the dynamics of retrieval during access to relational- namely temporal order- information from working memory. Critically, this approach enabled assessing WMC related changes in controlled memory retrieval without directly manipulating the presence of distractors or interference in the retrieval context.

A widely used task to measure temporal order memory is the judgments of recency (JOR) paradigm (Hacker, 1980; Liu, Chan, & Caplan, 2014; McElree & Dosher, 1993; Muter, 1979; Öztekin et al., 2008) in which participants are presented with a study list and asked to judge the relative recency of two test probes (e.g., which item appeared later in the study list). This task requires serial memory search operations, the efficiency of which would depend on both the maintenance of the items in an active state and executing the successive retrieval of items in order. Consequently, studying access to temporal order information from working memory can provide further insight with respect to how WMC modulates controlled retrieval.

Earlier work investigating the factors that modulate performance in the JOR task suggested that participants make strength-based judgments (Yntema & Trask, 1963). According to this hypothesis, the more recent probes evoke more strength than the less recent probes. Therefore, by making a strength comparison, it is possible to decide which item was presented more recently in the study list. In this case, the distance between the study positions of the test probes would determine the memory performance. The more distant the tested items were to each other, the better the memory performance would be (i.e., faster response times or higher accuracy). For instance, judging the recency of the 5th and 4th item in the study list would be easier than judging the recency of the 5th and 1st item as the strength level of the two items would be more similar in the former than the latter. As a consequence, strength based models predict performance to be modulated by the distance between the earlier and later probes. This pattern, however, was not observed with further investigations of the recovery of temporal order information (Hacker, 1980; Hockley, 1984; McElree & Dosher, 1993; Muter, 1979; Öztekin et al., 2008). What affected the memory performance was not the distance between the tested items but the recency of the later probe- the test probe which was presented later in the study list. Memory performance significantly increased (decreased RTs and increased accuracy) when the later probe was drawn from more recent positions in the study list, while earlier probe did not have an effect on the performance. These findings implicated that participants retrieved temporal order information via a serial search/scan through study list items.

A serial search through memory representations would operate as follows; participants start searching the studied items from the first (forward scan) or last (backwards scan) item in serial order, and the search is terminated upon reaching the later probe. For instance, in a backwards scan, the search will start with the last studied item, and the duration of the scan would depend on the study position of the later probe. The more recent the later probe, the shorter the scan would last. McElree and Dosher (1993) tested the serial memory search/scan hypothesis by employing the speed-accuracy trade-off (SAT) procedure to the JOR task with a 6-item study list (SAT procedure is explained in detail in the section below).

When all combinations of all-pairwise study positions (e.g., 2-1, 3-1, 4-1, 4-2, 4-3, 5-1, 5-2, etc.) were fitted with SAT retrieval functions, the observed pattern was in support for a serial scan process. Asymptotic accuracy, and retrieval speed (with more drastic differences in the intercept parameter) increased as the later probe was from the more recent positions. Accordingly, it has been suggested that the cognitive strategy that is used to recover temporal order information was a self-terminating backwards serial scan.

Present Study

In the current investigation, we tested HSs and LSs with a JOR task, using the SAT procedure. The aim of the study was to reveal whether HSs and LSs differ at recovering temporal order information, which typically requires an effortful controlled search through memory representations. To begin with, HSs and LSs might select different strategies whilst recovering temporal order information, which would be reflected in the retrieval dynamics measurements. Strategy selection has been suggested as a partial explanation to performance differences between HS and LS individuals (Cokely, Kelley, & Gilchrist, 2006; McNamara & Scott, 2001; Schelble, Therriault, & Miller, 2011). If HSs and LSs use different retrieval strategies we might observe differences in the retrieval dynamics across test probes as a function of study position; i.e., LSs using the distance between the items as a base for their judgments (the accuracy and retrieval speed will change as a function of the distance between the probes) and HSs using self-terminating serial scan (the performance will depend on the recency of the later probe). Alternatively, HSs and LSs might adapt the same strategy but LSs might not implement it as efficiently as HSs. If LSs were less efficient than HSs during the implementation of the strategy, the accessibility of temporal order information would be slower, which can be tracked by the retrieval speed estimates in the SAT functions, namely the rate of information accrual or the intercept, the time when information first becomes available. This might be due to the inability to generate and effectively use cues to delimit the search to relevant items. In particular, temporal order memory retrieval requires a serial search that demands more controlled processing than recognition memory judgments. As such, the time-course analysis of recovery of recency judgments can provide us more detailed information for the generalizability of variations in controlled processing as a function of WMC.

SAT Procedure

SAT is a variation of a deadline method in which subjects are signaled to respond at variable intervals following the onset of each test item, allowing a time course function that measures the growth of retrieval as a function of processing time. An important advantage of SAT over traditional paradigms is that it provides conjoint measures of the accuracy and speed of processing, independent of each other. This is in contrast to response time measures derived from traditional tasks, which cannot provide pure measures of processing speed because they are subject to speed–accuracy trade-offs (McElree, 2006). The SAT procedure can be used to measure the accuracy and speed of processing in a wide range of cognitive processes, including sentence comprehension (Foraker & McElree, 2007; Martin & McElree, 2009; McElree, Foraker, & Dyer, 2003), visual attention (e.g. Carrasco, McElree, Denisova, & Giordano, 2003; McElree & Carrasco, 1999), and memory (reviewed in McElree, 2006). Application of SAT in the memory domain has largely focused on investigations of item recognition (e.g. Benjamin & Bjork, 2000; Hintzman & Curran, 1994; McElree & Dosher, 1989; Öztekin & McElree, 2007; Wickelgren, Corbett, & Dosher, 1980), although it has been implemented to characterize relational memory processes as well (e.g. spatial order; Gronlund, Edwards, & Ohrt, 1997; n-back discriminations; McElree, 2001; and temporal order; McElree & Dosher, 1993).

Sampling the full time-course of retrieval also allows independently probing automatic versus controlled operations, as the output of automatic operations have typically been observed to be available before the output of controlled operations across a wide range of tasks (Hintzman & Curran, 1994; McElree, Dolan, & Jacoby, 1999; Öztekin & McElree, 2007, 2010). Accordingly, the SAT procedure enables independent estimation of both the timing and magnitude of the output of these processes via quantitative modeling routines (see Fig. 1.1 for illustration and description of a hypothetical SAT function).

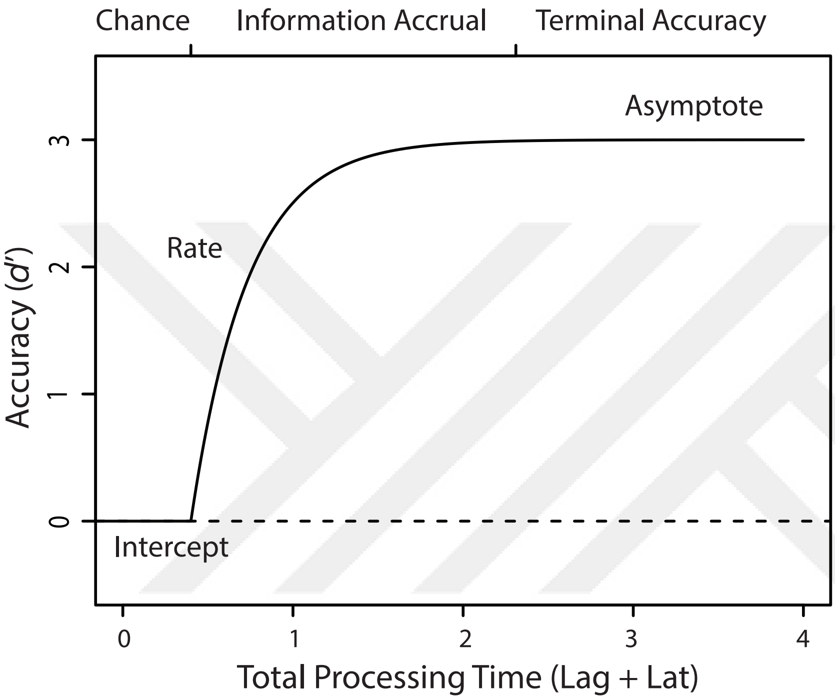

Figure 1.1. Illustration of a hypothetical SAT function that shows how accuracy (in d' units) grows over processing time (in seconds). The SAT curve reflects three phases: A period where performance is at chance (the departing point in time from chance is marked by the intercept parameter,), followed by a period of information accrual (the rise of this information accumulation is reflected by the rate parameter of the SAT function), and following this period, the maximum level of accuracy is reached, where performance does not improve any more (the asymptote parameter of the SAT function).

1.3. Method

Participants

Five-hundred and ninety adults were screened using the automated operation span task (Unsworth, Heitz, Schrock, & Engle, 2005) to attain WMC measures. 12 High Span (HS) individuals (upper quartile of the sample) and 12 Low Span individuals (lower quartile of the sample) participated in the experiment. Data from one participant of Low Span group, who failed to comply with the SAT procedure, was excluded from analyses, leaving 11 participants for the Low Span (LS) group. For the screening session, all participants received credit for Introduction to Psychology class via the Koç University subject pool system. For the experimental sessions, participants were compensated for their time.

Design and Stimuli

Operation span task. Participants were asked to solve math operations while trying to maintain a set of letters (F, H, J, K, L, N, P, R, S, T, Y) in their working memory. After each math operation, a letter was presented on the screen for 1000 ms. The list-length was varied from 3 to 7. At the test phase, participants recalled the order of the presented letters by marking the letters with numbers. Before the actual testing, participants were trained with three practice sets of list-length two. WMC scores were achieved by calculating the proportion of correct items marked at the correct position (Unsworth et al., 2005).

Judgment of recency task. The experiment was an adapted version of the Judgment of Recency task with the response deadline method. It consisted of six 50-min sessions, completed over a period of several weeks. In each session, there were 4 blocks of 140 experimental trials. Stimuli consisted of 18 consonants (b, c, d, f, g, h, j, k, l, m, n, p, r, s, t, v, y, z) displayed in lower case. Each study list comprised 5 consonants drawn randomly without replacement from the stimulus pool that had not appeared in the two preceding lists. Following a brief mask, participants were cued to respond to the test probe presented in upper case that consisted of two letters from the study list. Participants were asked to choose the most recently studied item. The position of the correct answer (left or right) was counterbalanced for each experimental session and participant.

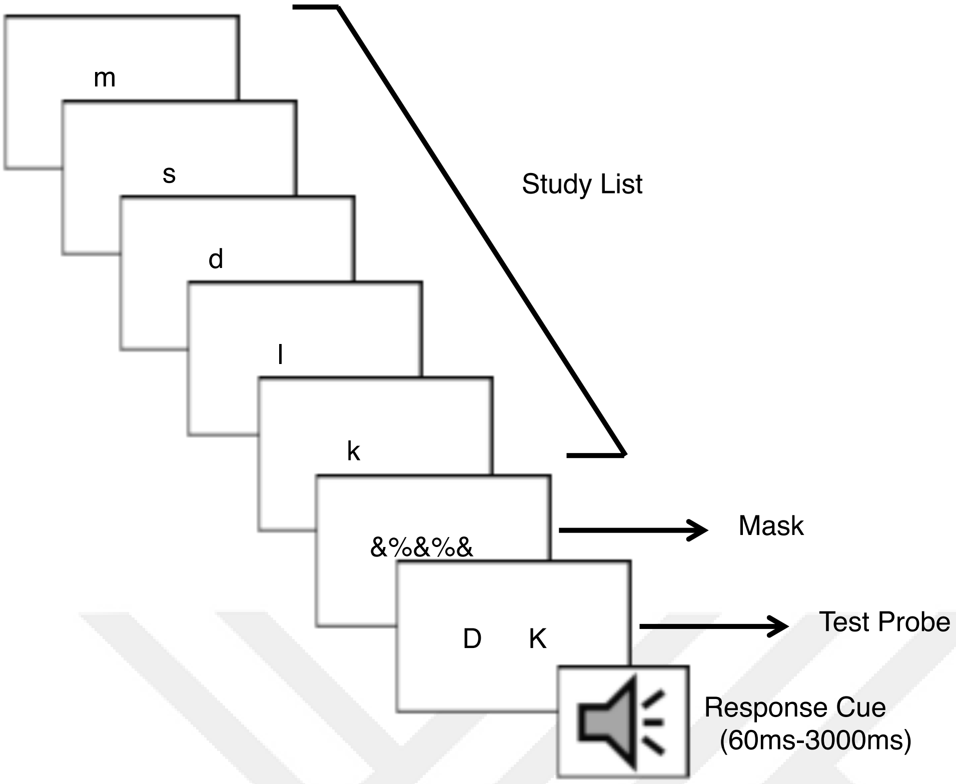

Figure 1.2. A sample trial sequence from the experimental procedure. Test probe consists of two items from the study list. K is the later probe, D is the earlier probe.

Procedure

Fig. 1.2 illustrates the sequence of a trial. Each trial began with a fixation point presented for 500ms. Each study item (a lowercase consonant) was presented one at a time for 500ms. Following the 5-item study list, a visual mask was presented for 500ms. Then, the test probes (two items from the study list presented as uppercase) appeared on the screen for the duration of the response deadline. At 60, 200, 300, 500, 800, 1,500, or 3,000ms after the onset of the recognition probe, a 50ms tone sounded to cue the participants to respond. Participants chose one of the presented items that corresponded to the “later item”, the item that appeared most recently in the study list. Participants indicated their response as quickly as possible after the onset of the tone by pressing a key. After indicating their response, participants were given feedback on their latency to respond. Participants were trained to respond within 300ms of the tone. They were informed that responses longer than 300ms were too slow and responses under 100ms were anticipations, and that both should be avoided.

There were ten conditions which utilized all possible study position pairings as the probe combinations: 2-1, 3-1, 3-2, 4-1, 4-2, 4-3, 5-1, 5-2, 5-3, and 5-4.

1.4. Results

Overview of Results

We first analyzed the data to reveal which strategy HSs and LSs applied to recover temporal order information. To do so, we assessed whether the study position of the earlier probe (SP-E), the probe that was presented earlier in the study list, or the study position of the later probe (SP-L), the probe that was presented later in the study list impacted the retrieval success and retrieval speed.

Next, we compared the efficiency of the groups while applying the chosen strategy (Also see SM-text for group differences in asymptotic accuracy measurements). We conducted between group comparisons on the performance measurements reflecting the availability (asymptotic accuracy, asymptote parameter estimates) and accessibility (rate and intercept estimates) of access to temporal order information. We also showed how WMC affects the retrieval dynamics when retrieval requires an effortful/controlled memory search.

Retrieval Dynamics

In order to obtain (equal-variance Gaussian) *d*' measures, an asymmetric *d*' scaling was calculated as such; $d' = [z\,(1|1) - z\,(1|2)] / 2^{1/2}$. Z here corresponds to the standard normal deviate of the probability of responding that the most recent item was the first alternative, given that the test probe was either the first (1|1) or the second (1|2) alternative. We estimated the retrieval dynamics by fitting the individual participants' data and the average data (derived by averaging *d*' values for each condition across participants) with an exponential approach to a limit:

$$d'\ (t) = \lambda\left(1 - e^{-\beta(t-\delta)}\right), t\ > \delta, \text{else } 0. \qquad (1)$$

In Equation (1.1), $d'(t)$ is the predicted d' at time t; λ is the asymptotic accuracy level reflecting the overall probability of recognition; δ is the intercept reflecting the discrete point in time when accuracy departs from chance ($d' = 0$); β is the rate parameter, which indexes the speed at which accuracy grows from chance to asymptote. Previous studies have indicated that this equation provides a good quantitative summary of the shape of the SAT functions (Dosher, 1981; McElree, 2001; McElree & Dosher, 1989; Wickelgren & Corbett, 1977; Wickelgren et al., 1980).

The quality of the model fits were assessed by: (a) The value of an adjusted-R^2 statistic (Reed, 1973); (b) the consistency of parameter estimates across participants; (c) evaluation of whether the fit yielded systematic deviations that could be accounted for by additional parameters. These latter two metrics were assessed by statistical tests conducted on the parameter estimates derived from the model fits across participants.

Initially, we fit the full model (10λ-10β-10δ) to individual participants' data for both groups to examine the impact of WMC on each test probe combination and to evaluate the patterns in support for the serial search mechanism.[1] Parameter estimates derived from the full model were used for the statistical analysis of WMC and serial scan effects (see Tables SM-1.1, SM-1.2, SM-1.3 and SM-1.4 in Supplemental Material 1 for a complete list of the parameter estimates from the full model and adjusted-R^2 values). We additionally tested models by varying the number of parameters allocated to different conditions in order to attain the best fitting model.

Strategy choice effects on retrieval dynamics

As overviewed in the Introduction, there are two possible strategies participants could employ in the JOR task. If participants judge the recency of the items by making a strength comparison, we expect performance to improve as a function of the distance between SP-E

---

[1] The full model estimated one asymptote (λ), one rate (β) and one intercept (δ) parameter for each condition.

and SP-L. If, on the other hand, participants apply a self-terminating serial scan, performance would be determined by the changes in the SP-L alone.

Probe distance effects

It is possible to assess distance effects by comparing conditions in which SP-L is held constant and only SP-E is varied. We started with holding SP-L 5 constant and varying SP-E; a 2 (Group [HS vs. LS]) x 4 (SP-E : 1, 2, 3, 4 compared with SP-L 5) ANOVA indicated that SP-E did not have a measurable effect on the asymptote (p = .54], the rate (*p* = .40), or the intercept (*p* = .69). We did not observe interactions between the WMC and the SP-E on asymptote (*p* = .56), rate (*p* = .30), and intercept parameters (*p* = .50).

We next analyzed variations in SP-E for SP-L 4. A 2 (Group [HS vs. LS]) x 3 (SP-E: 1, 2, 3 compared with SP-L 4) mixed ANOVA analysis showed no main effect of SP-E on asymptote (*p* = .31), rate (*p* = .24), and intercept (*p* = .28) parameters. There was also no Group x SP-E interaction for asymptote: (*p* = .86), rate (*p* = .75), and intercept: (*p* = .21) measurements.

Finally, a 2 (Group [HS vs. LS]) x 2 (SP-E: 1, 2 compared with SP-L 3) revealed that while SP-E had no main effect on asymptote (*p* = .11) and rate measures (*p* = .21), it significantly affected the intercept parameter [$F(1,21) = 10.66, p < .005, \eta_g^2 = .19$]. Similar to the findings above, there was no interaction of SP-E with WMC for all parameter estimates (asymptote: *p* = .49, rate: *p* = .66, intercept: *p* = .21). Further post hoc tests on the intercept measures showed that the SP-E effect was only apparent for LSs with 3-1 test probe combination having significantly lower intercept (*t*(21)=-3.462, p < .003) measures than 3-2 combination. We speculate that LSs might have applied a different strategy on certain trials while recovering the 3-1 combination.

According to temporal context model (Howard & Kahana, 2002), nearby positions are coded in a similar fashion so that judging the nearby positions would be more difficult than judging the positions that are not nearby. We did not observe such differences for SP-L 4 or SP-L 5. This might be because these items are still available in memory and the backward self-terminating serial scan can be applied when the item is still available (McElree, 2006). However, on trials when the item availability was poor, they might have switched their strategy to a strength-based comparison. A similar pattern was observed in a previously tested JOR paradigm (see Klein, Shiffrin, & Criss, 2007), participants substituted their strategy by a strength comparison when contextual based judgments did not work. This would explain the performance differences between probe combinations 3-1 and 3-2 and also the lack of WMC impact on the intercept parameter for probe 3-1. We did further analysis to investigate whether this might be due to a primacy effect (See SM-Text), however, we did not observe any primacy effects. Specifically, when SP-E was the 1st study list position, memory performance was not better compared to other positions of SP-E. This is also consistent with previous studies that also showed no evidence for primacy effects in JOR paradigm (e.g., Klein, Criss, & Shiffrin, 2004).

Self-terminating serial scan

After establishing that the SP-E did not modulate the retrieval dynamics, we assessed the impact of the SP-L. With a similar approach, this time we held the SP-E constant across conditions and investigated how SP-L modulated memory performance. A 2 (Group [HS vs. LS]) x 4 (SP-E: 1, SP-L: 2, 3, 4, 5) mixed ANOVA analysis on the asymptote [$F(2.14,44.92) = 26.37, p < .001; \eta_g^2 = .33$][2], rate [$F(2.59,54.33) = 8.17, p < .001; \eta_g^2 = .19$], and intercept [$F(1.55,32.47) = 24.62, p < .001; \eta_g^2 = .47$] parameters showed that SP-L had a significant impact on the retrieval accuracy and retrieval speed. The more recent the SP-L, the better the performance was for both groups. Group by condition interactions on asymptote

---

[2] Throughout the manuscript degrees of freedom (df) are Greenhouse-Geisser corrected in ANOVAs for repeated-measures factors with more than two levels.

[*F*(2.14,44.92) = 4.58, $p < .02$, $\eta_g^2 = .08$] and intercept parameters [*F*(1.55,32.47) = 3.19, $p < .07$, $\eta_g^2 = .10$] showed that LSs' performance were affected more with the changes in SP-L compared to HSs. Namely, LSs performed worse than HSs when SP-L was drawn from less recent study positions.

A 2 (Group [HS vs. LS]) x 3 (SP-E: 2 compared with SP-L: 3, 4, 5) mixed ANOVA analysis conducted on asymptote [*F*(1.90, 39.83) = 36.56, $p = .001$, $\eta_g^2 = .35$], rate [*F*(1.76, 36.88) = 6.83, $p < .005$, $\eta_g^2 = .18$], and intercept [*F*(1.41, 29.67) = 27.51, $p < .001$, $\eta_g^2 = .39$] parameters also indicated a significant main effect of SP-L. We also observed significant Group x SP-L interactions on both asymptote [*F*(1.90, 39.83) = 12.05, $p < .001$, $\eta_g^2 = .15$] and intercept parameters [*F*(1.41, 29.67) =5.18, $p < .02$, $\eta_g^2 = .11$] indicating a similar pattern explained above.

Finally, a 2 (Group [HS vs. LS]) x 2 (SP-E: 3 compared with SP-L: 4, 5) revealed that SP-L determined the changes in asymptote [*F*(1, 21) = 27.66, $p < .001$, $\eta_g^2 = .24$], rate [*F*(1, 21) = 39.76, $p < .005$, $\eta_g^2 = .20$], and intercept parameters [*F*(1,21) = 6.72, $p < .03$, $\eta_g^2 = .13$]. We observed Group x SP-L interaction only for asymptote measurements for this analysis [*F*(1, 21) = 13.18, $p < .003$, $\eta_g^2 = .13$].

After establishing that SP-L modulated the memory performance, we collapsed SP-E across SP-L and conducted a 2 (Group [HS vs. LS]) x 4 (SP-L: 2, 3, 4, 5) mixed ANOVA analysis on the asymptote [*F*(1.82,38.31) = 31.22, $p < .001$; $\eta_g^2 = 0.36$], rate [*F*(2.51,52.70) = 25.06, $p < .001$; $\eta_g^2 = .2$], and intercept [*F*(1.38,29) = 23, $p < .0001$; $\eta_g^2 = .45$] parameters, which showed that SP-L had a significant impact on the retrieval accuracy and retrieval speed. Specifically, HSs and LSs were more accurate and faster when SP-L were from more recent study positions. WMC also had a significant main effect on both retrieval accuracy and speed measurements: asymptote [$F(1,21) = 10.8$, $p < .005$; $\eta_g^2 = .24$], rate [$F(1,21) = 4.64$, $p < .05$; $\eta_g^2 = 0.06$], intercept [$F(1,21) = 17.54$, $p < .0005$; $\eta_g^2 = .17$]. There was also interactions between group and SP-L on asymptote [$F(1.82,38.31) = 6.29$, $p < .006$; $\eta_g^2 = .10$] and intercept $F(1.38,29) = 3.33$, $p < .07$; $\eta_g^2 = .11$] measurements showing a more profound impact of SP-L on LSs' memory performance. These effects are further investigated in the following sections.

There were significant linear trends for both groups showing that asymptote estimates increased as a function of SP-L [HS: estimated slope = 2.79, $t(63) = 4.956$, $p < .001$, LS: estimated slope = 4.99, $t(63) = 8.48$, $p < .001$]. For the rate parameter, the pattern was similar [HS: estimated slope = 13.39, $t(63) = 2.26$, $p < .03$, LS: estimated slope = 12.39, $t(63) = 2.008$, $p < .05$]. There was also an operative linear trend for the intercept parameter [HS: estimated slope = -1.58, $t(63) = -3.589$, $p < .001$, LS: estimated slope = -3.58, $t(63) = -7.77$, $p < .0001$]. Specifically, the intercept parameter was faster as a function of SP-L. Accordingly, our analyses indicated that performance varied as a monotonic function of SP-L, indicating a backwards serial search strategy.

The most important indicator of a serial scan mechanism is the delays in the initial availability of the temporal order information (e.g., see McElree & Dosher, 1993). Depending on the chosen strategy, forward or backwards, the intercept parameter will vary as a function of the recency of the later probe. In our case, retrieving temporal order information for SP-L 2 comparisons took longer than all the other combinations (slowest intercept parameter for each group), while significantly improving as a function of SP-L, which is suggestive for backwards scanning strategy. The intercept measurements allocated for SP-L 5 was significantly faster than all other conditions, suggesting that participants were much faster while accessing the temporal order information when the test probe contained the most recently studied item.

To further investigate the serial scan effects on the data, we sought to identify models that most adequately describes the data taking both model fit and flexibility account (see SM-Text for the detailed model selection procedure). We started fitting the data with a null model, which allocated a common asymptote [λ], a common rate [β], and a common intercept [δ] parameter to each probe combination. We then tested models, which varied the parameters as a function of the later probe.

Allocating different parameter values to different later probe conditions for both retrieval accuracy (asymptote), and retrieval speed (intercept and rate) parameters as a function of the later probe significantly increased the adjusted-R^2 statistics for both groups (see SM-text). Varying the parameters as a function of the SP-L helped us to explain the systematic deviations, and revealed the linear trends that were also prominent in the parameter estimates derived from the full models[3]. The observed linear trends, the significant impact of SP-L on the parameter estimates, and the properties of best fitting models evidently show us both HSs and LSs applied the self-terminating serial strategy. We next investigated whether there were any differences between the two groups whilst applying the strategy.

WMC Impact on Temporal Order Memory Retrieval

---

[3] We did not use the parameter values estimated by the most adequate model to statistically compare the two groups performance. HS and LS groups' empirical data were best explained by models with different numbers of parameters, hence it would not be appropriate to perform a between group comparison on the estimated parameters from these models. However, we used the parameter estimates from the most adequate models for each group to depict the SAT functions in Fig. S3. A more detailed description is provided in SM-text.

After establishing the data can be explained by a backwards serial scan, we evaluated WMC related effects on the asymptote and retrieval dynamics parameters.

Retrieval Accuracy

Results from independent samples t-tests and the average of the parameter estimates with the standard deviations from both groups are shown in Table 1.1 Between-group comparisons conducted on the asymptote parameter estimates derived from the SAT functions indicated a lower asymptote for LS compared with HS group for all conditions, except for those containing SP-L 5. This pattern also holds for the empirical measurements of asymptotic accuracy (See SM-Text 1.1 and 1.2.) Thus, WMC impacted the probability of successful retrieval of temporal order information unless the later probe is the latest presented item, which is presumably still in the FoA.

Retrieval Speed

As shown in Table 1.2, the groups did not differ in the rate parameter estimates in any of the test probe combinations. However, the groups had significant differences in the intercept parameter estimates. Table 1.3 shows the between group comparisons on the intercept estimates for each test probe combination. Overall, the intercept parameter, which marks the point in time when information first becomes available in memory, was slower for LSs compared to HSs. The observed differences between the groups were apparent in all test probes, except for SP-L 5 (i.e. SP-L 4, SP-L 3, and SP-L 2). In contrast to rate estimates, WMC had an obvious impact on the intercept measurement except for SP-L 5 (See the depiction of the SAT functions from the full fit in Fig. 1.3). This finding is consistent with previous research indicating the most recently studied item is maintained in the current focus of attention (FoA) hence does not require an effortful search through memory representations (McElree, 2006; Mızrak & Öztekin, 2016; Öztekin, Gungor, & Badre, 2012; Öztekin & McElree, 2007, 2010). Accordingly, that WMC effects on retrieval speed and accuracy were prominent when controlled processing is required, and not present when the memory judgment entailed matching the probe to the contents of focal attention is in line with the contention that WMC selective impacts controlled processing.

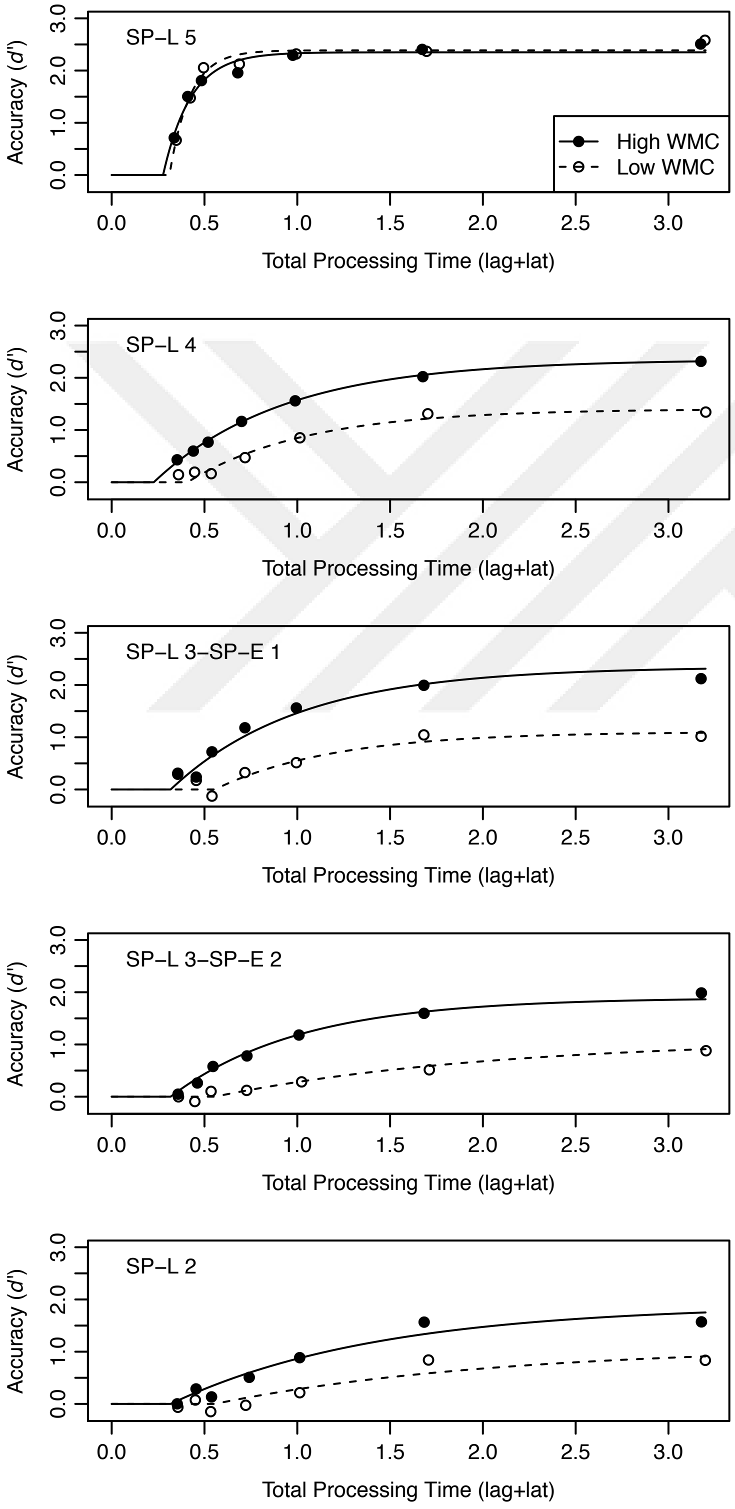

Figure 1.3. Accuracy (in *d'* units) plotted for test probe combinations as a function of total processing time (in seconds) for the average high-span (HS) and low-span (LS) groups. For each test probe combination, except from the SP-L 5 combinations, LSs had later intercepts and lower asymptote levels derived from the SAT function parameters. The symbols indicate empirical data points, and the smooth curves indicate the model predictions derived from Equation 1.1. SP-L 5 = all test probes in which the later probe was the 5th item in the study list. SP-L 4 = all test probes in which the later probe was the 4th item in the study list. 3-1 = test probe which had the 3rd item as the later probe and 1st item as the earlier probe from the study list. 3-2 = test probe which had the 3rd as the later probe and 2nd item as the earlier probe from the study list. SP-L 2= test probe which has the 2nd item from the study list as the later probe.

As described in the previous sections, there were significant linear trends in estimated parameters for both groups, which is consistent with the application of a backwards serial scan. Both groups had slower intercept parameters when SP-L was drawn from earlier study list positions. What is striking here is that the group differences were not limited to the intercept estimates across SP-L conditions: the data also showed a WMC effect on the slope of the linear increase in the intercept parameters across SP-L conditions. In other words, LS individuals further slowed down with SP-L being drawn from earlier study positions [$t(63) = 3.132$, $p < .003$]. This finding implicates that LS group initiated the serial scan strategy later than HS group, and they were also less efficient in completing the search after the search started.

As mentioned above, the variation in the intercept parameter of the SAT function is seen as a major index of the application of the serial scan mechanism (McElree & Dosher, 1993). In our case, both HSs and LSs employed a backwards serial scan strategy to recover temporal order information. However, LSs intercept measurements were significantly slower than HSs, showing that LSs started serial memory search operations later than HSs. The data further suggested that LSs were less efficient in applying the backwards scanning strategy. That WMC effects were prominent on the intercept parameters of the SAT function strongly indicates the group differences to arise from efficacy of the serial memory search.

Summary of Results

In summary, WMC had a significant impact on both retrieval accuracy, measured by empirical asymptotic accuracy measurements and asymptote parameters from the SAT functions, and retrieval speed, namely on the intercept parameter. Specifically, high-span group outperformed low-span group by having higher availability of the temporal order information and by accessing the required information faster. The data further implicated the WMC related differences in retrieval dynamics to specifically reflect less efficient serial memory search. However, this pattern was absent when the more recent probe of the test probes were the last item in the study list, which can be maintained in the current focus of attention. That no group differences are observed for the contents of FoA supports previous research suggesting that WMC differences emerge only under certain conditions (Barrett et al., 2004; see Engle, 2002 for a review; Öztekin & McElree, 2010), namely, in situations when the information can only be accessed via strategic, controlled operations.

1.5. Discussion

In this study, we investigated the role of WMC on the dynamics of temporal order memory retrieval. Specifically, we tested HS and LS individuals with a relative JOR paradigm, in which they were asked to retrieve the temporal information of the items presented and judge the recency of the test probes. To independently assess retrieval accuracy and retrieval speed, we employed the response-deadline SAT procedure and obtained full time-course retrieval functions. The overall accuracy of temporal order information retrieval, measured by asymptotic accuracy (average of the two last response deadlines), was lower for LSs than HSs except for the test probes containing the most recently studied item. This lower performance was also evident in the asymptote estimates derived from quantitative modeling of the SAT functions. Consistent with the previous research showing that LSs perform worse than HSs across a variety of cognitive tasks (reviewed in Engle, 2002, 2010), WMC measures correlating with the probability of retrieval is not surprising. The more novel contribution of the current investigation concerns the dramatic retrieval speed differences across the two groups during access to temporal order information in working memory: LSs were slower in engaging in the controlled serial memory search operations that access temporal order information from working memory. In particular, our data pinpointed the WMC group differences to predominantly affect the intercept parameters of the SAT function across successive retrieval operations. Below, we provide possible explanations to the observed findings.

WMC selectively impacts controlled retrieval

It is noteworthy to first emphasize that WMC did not have an impact on the success or speed of access to temporal order information that was still in the focus of attention, a case when no memory operation is required, and hence retrieval can be regarded as largely automatic. Theories explaining the underlying mechanisms that lead to individual differences in WMC capitalize on controlled attention with LSs performing worse in situations that necessitate controlled processes (Engle & Kane, 2004), with no measurable differences in automatic processing. While shunting information outside of the focal attention requires controlled retrieval, the contents of focal attention exhibit privileged access (McElree, 2001, 2006; McElree & Dosher, 1989; Mızrak & Öztekin, 2016; Öztekin et al., 2012; Öztekin & McElree, 2007, 2010). The lack of group differences in either the success or speed of retrieval in the current study supports the contention that WMC selectively affects controlled processing, rather than a leading to a more global deficit (Redick et al., 2012). In contrast, for material outside of current focus of attention, WMC impacted on both the success and efficiency of the serial search operations deployed to access temporal order memory.

Serial memory search: WMC does not affect the strategy to recover temporal order memory retrieval

Earlier theories (Yntema & Trask, 1963) of judgments of recency suggested that the distance between the test probes should have an effect on the temporal order retrieval performance; however, we did not observe this distance effect in our study. Alternatively, the recency of the more recent probe in the test probes (the later probe) mediated the retrieval success and retrieval speed for both HSs and LSs. The performance decreased linearly as the more recent probe was drawn from earlier study positions and this decrease was more prominent for LSs compared to HSs. In this regard, our findings were consistent with previous literature in showing that access to temporal order information in the JOR task requires a serial search through memory representations (Hacker, 1980; McElree, 2006; McElree & Dosher, 1993; Öztekin et al., 2008). Notably, the dynamics of memory retrieval obtained from quantitative modeling applied to each individual participant's, as well as the group data implicated a serial search strategy for both HS and LS groups. This pattern was extracted from faster retrieval dynamics (earlier intercepts) and higher retrieval success when the more recent test probe was from later study positions. McElree and Dosher (1993) showed that the recency of the later probe had drastic impact on the intercept parameter of the SAT function, with more recent probes having earlier intercepts. They interpreted this pattern as indicative of a backward serial search (see also Hacker, 1980), that starts from the last item in the study list and if the tested probe is the last item, this search will be less effortful and take less time. Our findings indicated a similar pattern for both groups: When the later probe matched the last item in the study list; both groups exhibited the fastest intercept and highest accuracy rate amongst all the test probes. Moreover, changes in the intercept parameter as a function of the position of the later probe further confirmed that participants were engaging in a serial search, namely the intercept was earlier as the position of the later probe was more recent. We next discuss WMC's impact on the efficiency of this serial memory search.

WMC and efficiency of serial memory search

Although HSs and LSs implemented the same retrieval strategy, namely a serial memory search, LSs were not as efficient as HSs. This was evident in both retrieval success and retrieval speed measures. Quantitative modeling of the SAT functions indicated that LS individuals were slower (reflected in the intercept parameter of the SAT function) in initiating the successive serial memory search operations that access temporal order information from working memory. This is in contrast to previous investigations assessing the relationship between WMC and item recognition (Öztekin & McElree, 2010), which indicate that LS individuals exhibit similar retrieval speed measures when accessing item information from working memory, in the absence of interference. Speed differences were only prominent when resolving interference, during which LSs were slower. Taken together, these data suggest that group differences with respect to retrieval speed measures emerge when controlled processing is required, either due to the presence of interference in the retrieval context or depending on the type of information (i.e. item versus relational) that needs to be accessed from memory. In the following sections, we discuss the observed group differences in more detail.

WMC effects on availability of temporal order memory. For successful retrieval, memory representations need to be available in memory. Availability of the memory representations may depend on representations being encoded effectively, and maintained in an active state until the time of retrieval. If the memory representations are not available for retrieval, this might be due to representations not being actively maintained in working memory or they become unavailable due to loss of strength or interference of task-irrelevant items. The asymptote measurements in our study reflect the availability of the target information required for the task. When given enough time, if participants cannot recover the temporal order information this can be presumably interpreted as a lack of availability of the temporal representations of the study items in memory. Our findings showed that LSs had lower accuracy than HSs, measured by empirical asymptotic accuracy and SAT asymptote estimates. This finding, complementing other studies in the literature (Kane, Bleckley, Conway, & Engle, 2001; Kane & Engle, 2002, 2003; see Unsworth & Engle, 2007 for a review), might indicate that LSs are worse than HSs at maintaining representations in an active state. In JOR paradigm, participants need to protect the contents of their memory while switching between multiple active representations, in order to reach the required information (the temporal position of the items) for the recency judgment. The present data offer two possible explanations to the lower availability of the temporal order information for the LS group. It is possible that LSs were not able to encode the study materials well in the first place. Encoding temporal order information may require more elaborative encoding, and HSs might better adjust themselves for efficiently encoding the temporal order of the study material. Alternatively, the two groups may not differ at sufficient encoding, however accuracy differences might arise solely due to the inefficiency of the memory search for the LS group. Previous research suggests that LSs are not as good as HSs at filtering out the irrelevant information and allocating their attention on the relevant material (Kane et al., 2001; Öztekin & McElree, 2010; Unsworth, Schrock, & Engle, 2004), and that they are prone to internal and external distraction. LSs' being less effective while they are allocating their attention is a well-studied phenomenon. Therefore, we believe the latter explanation is more likely.

WMC effects on accessibility of temporal order memory. One salient advantage provided by the SAT procedure is that it enables independent assessments of retrieval success and processing speed. This allowed us to differentiate the dynamics of the retrieval operations engaged to access temporal order memory, independent of differences in terminal accuracy. Our data showed that in addition to the lower availability of the temporal order information, LSs were slower in engaging in the successive serial memory search operations that access temporal order information. Retrieval dynamics measures obtained from the SAT procedure enable separately assessing both the time point in which information first becomes available (the intercept parameter) and the rate of the accumulation of information in memory (the rate parameter). Previous investigations of the time course of JOR (McElree & Dosher, 1993) have indicated serial memory search operations to reflect on the intercept parameter. In other words, each successive retrieval operation that accesses study items in their temporal order takes time, and delays the intercept parameter. In line with these previous findings, our data specifically pinpointed group differences to occur on the intercept parameter of the SAT function, implicating that the LS group is less efficient in deploying these serial memory search operations that access temporal order information. This slower retrieval of temporal order information could stem from (a) delayed initiation of the serial scan process, (b) slower scanning through the memory representations, or a combination of both. Indeed, our data implicates a combination of both a) and b). Below, we discuss these explanations in more length.

*Delayed initiation of the serial scan***.** Successfully judging the recency of the items requires activating representations outside the focus of attention, execution of the serial memory search, and switching attention to the successively retrieved representations during this controlled memory search. For this entire process to be successful, the activated item representations should be ordered by their position in the study list. If WMC reflects the ability to maintain context binding (Oberauer, 2005), one possibility is that LSs might not have maintained temporal context bindings as well as the HS group. This would slow down correctly reinstating the temporal positions of the studied items.

Studies examining the controlled search differences between HSs and LSs found similar findings, showing that LSs were not as efficient as HSs in the search process (e.g., Spillers & Unsworth, 2011; Unsworth, Spillers, & Brewer, 2012). It is possible that LSs are not as good as HSs when setting up an overall retrieval plan (Unsworth, Brewer, & Spillers, 2013). In JOR task, the retrieval plan might be to strongly associate the items with their positions that would make reinstating the context faster. That said, previous research suggested that the groups differ in generating efficient retrieval cues to search memory with (Spillers & Unsworth, 2011). Setting up a plan and generating cues to use during retrieval are important and LSs might be doing both of these but what affects their efficiency? For instance, they might also have a cue present at retrieval, but whether this cue leads them to the specified relevant item or not is an important component that would affect successful retrieval of the item.

WMC is measured by complex span tasks, which intrudes processing of the memoranda by presenting additional operations. Participants have to maintain the recently encountered representations in a state to retrieve them later. When the pattern of errors to this task was examined (Unsworth & Engle, 2006), it has been shown that LSs could not retrieve the items from earlier positions of the lists as much as the items that were presented later in the lists. LSs also had higher output transpositions (an item recalled correctly but in an incorrect position) than HSs. Accordingly, it has been suggested (see Chow & Conway, 2015; and Unsworth & Engle, 2006) that, LSs employ less efficient memory cues that disallows them to remember the items' positions correctly. The output transpositions occurred mostly for the items from middle positions in the lists, which suggests that temporal-contextual cues for these items were not diagnostic enough. Additionally, Spillers and Unsworth (2011) employed a delayed free recall paradigm to investigate how WMC impacted the use of internally generated temporal-contextual cues. They showed that, although LSs and HSs initiated the recall process similarly, LSs were not as efficient as HSs. While HSs benefited from using the recalled items as a cue to recall other items, LSs were unable to do this. These findings suggest that LSs are less efficient in using temporal-contextual cues during retrieval, which could explain the delay in initiating the serial memory search. However, the observed differences in the linear trends on the intercept parameters cannot be explained by just the delay in initiating the serial memory search. This latter finding further implicates that LSs were also less efficient in executing the serial scan operations after they had been initiated.

*Slower memory search***.** After the participants initiated the serial scan, what is left is to scan through the memory representations until the later probe matches one of the successively retrieved representations. In order to do this, participants need to carry out multiple retrieval operations, and compare the test probes with the elements in memory. Comparing activated memory representations with the probes, and judging whether they match or not might affect the retrieval speed. However, in a previous study that investigated time-course of WMC related effects in item recognition (Öztekin & McElree, 2010), there were no speed differences between the HS and LS groups when they judged whether the probe belonged to the study list or not. Therefore, this explanation is less likely. Consistent with this finding, our results showed that there were no differences between groups in the speed of information accumulation, reflected in the rate parameter. On the other hand, the difference between the groups in the linear trend in which the intercept parameter further slows as a function of the study position of the later probe might arise from differences in the speed of scanning through memory representations. This process requires switching the focus of attention from one representation to another at each successive retrieval operation, which would require controlled attention resources. We suggest that LSs serial memory search operations might be further delayed due to the necessity of attentional control at each successive retrieval operation during the serial scan. Accordingly, our findings implicate both that a) LSs were delayed in initiating the serial scan and b) were further less efficient in carrying out this controlled serial memory search.

1.6. Conclusion

In this study we evaluated the impact of working memory capacity on the recovery of temporal order information in retrieval accuracy and retrieval speed measures. Our results suggest that WMC predicts the ability to search through items in memory by slowing the serial search operation employed to access temporal order information. Although both groups applied the same strategy to recover temporal order information, namely a self-terminating backwards serial scan, low WMC group was slower and impaired in this search process. A serial search through items in memory is an exhaustive controlled search process that requires cognitive control. We showed that initiation of the serial memory search operations was delayed for individuals with low WMC compared to high WMC. Low WMC group was also less efficient in completing the serial scan. Accordingly, the data implicate that retrieval of temporal order information from working memory was slower for low WMC group due to both a delay in initiation of a backwards serial scan strategy, and a less efficient/slower serial memory search after the scan had started. These findings are consistent with the previously observed delayed controlled processes for low WMC in the face of interference/distractors, and extends this notion to other contexts that require controlled processing, such as recovery of relational information.

Table 1.1
Between group comparisons of asymptote estimates derived from the full model and the means for HS and LS groups. LP = Later probe.

Test Probe Combination	LP	HS- Mean(SD)	LS- Mean(SD)	*t*	df	*p*
2-1	2	1.64 (0.90)	0.93 (0.70)	2.12	20.58	0.047
3-1	3	2.14 (0.65)	1.21 (0.62)	3.49	20.948	0.002
3-2	3	2.01 (0.65)	0.91 (0.63)	3.78	20.346	0.001
4-1	4	2.50 (0.51)	1.55 (0.81)	3.28	16.774	0.004
4-2	4	2.37 (0.56)	1.38 (0.64)	3.94	19.997	0.001
4-3	4	2.29 (0.64)	1.42 (0.69)	3.13	20.387	0.005
5-1	5	2.51 (0.47)	2.47 (0.42)	0.19	20.976	0.849
5-2	5	2.50 (0.49)	2.46 (0.47)	0.19	20.921	0.846
5-3	5	2.49 (0.54)	2.49 (0.44)	-0.03	20.178	0.978
5-4	5	2.41 (0.52)	2.43 (0.40)	-0.08	20.399	0.935

Table 1.2
Between group comparisons of rate estimates derived from the full model and the means for HS and LS groups. LP = Later probe.

Test Probe Combination	LP	HS- Mean(SD)	LS- Mean(SD)	*t*	df	*p*
2-1	2	5.24 (3.68)	7.61 (5.18)	-1.34	17.90	0.199
3-1	3	5.19 (6.06)	5.96 (6.16)	-0.31	20.76	0.770
3-2	3	6.59 (6.59)	8.86 (5.72)	-0.88	20.95	0.388
4-1	4	5.76 (6.31)	7.94 (7.20)	-0.77	19.99	0.452
4-2	4	2.90 (3.98)	6.92 (6.32)	-1.81	16.62	0.089
4-3	4	3.11 (2.80)	5.77 (5.80)	-1.37	13.04	0.193
5-1	5	11.48 (6.70)	13.79 (6.70)	-0.08	20.82	0.419
5-2	5	12.37 (7.43)	11.17 (6.78)	0.40	21.00	0.689
5-3	5	10.11 (7.94)	10.71 (7.07)	-0.19	20.99	0.850
5-4	5	7.50 (7.76)	12.62 (8.88)	-1.50	20.00	0.158

Table 1.3

Between group comparisons of intercept estimates derived from the full model and the means for HS and LS groups. LP = Later probe.

Test Probe Combination		HS- Mean (SD)	LS- Mean (SD)	*t*	df	*p*
2-1	2	0.759 (0.430)	1.380 (0.730)	-1.47	15.97	0.024
3-1	3	0.418 (0.170)	0.593 (0.299)	-1.69	15.78	0.109
3-2	3	0.570 (0.290)	1.090 (0.561)	-2.75	14.71	0.015
4-1	4	0.292 (0.080)	0.470 (0.291)	-1.94	11.53	0.077
4-2	4	0.293 (0.240)	0.603 (0.215)	-3.25	20.99	0.003
4-3	4	0.410 (0.280)	0.520 (0.296)	-0.88	20.60	0.380
5-1	5	0.255 (0.100)	0.298 (0.090)	-1.06	20.99	0.299
5-2	5	0.266 (0.162)	0.301 (0.124)	-0.58	20.37	0.570
5-3	5	0.313 (0.122)	0.295 (0.068)	0.44	17.46	0.660
5-4	5	0.305 (0.129)	0.289 (0.110)	0.31	20.99	0.762

CHAPTER II

Relationship between Emotion and Forgetting

2.1. Abstract

A major determinant of forgetting in memory is the presence of interference in the retrieval context. Previous research has shown that proactive interference has less impact for emotional compared to neutral study material (Levens & Phelps, 2008). However, it is unclear how emotional content affects the impact of interference in memory. Emotional content could directly affect the build-up of interference, leading to reduced levels of interference. Alternatively, emotional content could affect the controlled processes that resolve interference. The present study employed the response deadline speed-accuracy trade-off (SAT) procedure to independently test these hypotheses. Participants studied 3-item lists consisting of emotional or neutral images, immediately followed by a recognition probe. Results indicated a slower rate of accrual for interfering material (lures from previous study list) and lower levels of interference for emotional compared to neutral stimuli, suggesting a direct impact of emotion on the build up of interference. In contrast to this beneficiary effect, the resolution of interference for emotional material was less effective than neutral material. These findings can provide an insight into the interactions of emotion and memory processes.

2.2. Introduction

Memory enhancement for emotional material is a well-studied phenomenon (see Buchanan, 2007; Doerksen & Shimamura, 2001; Hamann, 2001 for reviews). Research on emotion-memory interactions has suggested that emotional information may be more likely to be remembered and persist longer in memory than neutral material (Buchanan, Etzel, Adolphs, & Tranel, 2006). Emotional content has influence on the quantity and quality of events remembered as well as the total amount of correctly remembered details of the events (Kensinger & Schacter, 2008). Additionally, emotional memories are remembered with a greater sense of recollective experience (Sharot & Yonelinas, 2008) that is strengthened over time (Ritchey, Dolcos, & Cabeza, 2008).

Although effects of emotional content on remembering are well documented, the exact nature of this relationship is not as clear. Here we attempted to provide a better understanding of the interactions between emotion and memory processing using an in depth investigation of the impact of emotion on forgetting. A mechanistic explanation of the relationship between emotion and forgetting could provide insight into the underlying mechanisms that modulate the interactions between emotion and memory processes.

A major cause of forgetting over the short and the long term is the presence of interference in the retrieval context. When faced with proactive interference (PI), pre-learned yet irrelevant material interferes with the subsequent encoding and/or retrieval of newer information (Gardiner, 1972; Keppel & Underwood, 1962; Watkins & Watkins, 1975; Wickens, 1970). Because PI affects the efficient use of working memory (WM) resources, it can drastically limit ongoing cognitive processes. Hence control of PI is a critical determinant of successful performance for many cognitive tasks that depend on WM resources. Although PI is a well-studied phenomenon, the majority of the literature has focused on the effects of PI with nonemotional stimuli (Badre & Wagner, 2005; D'Esposito, Postle, Jonides, & Smith, 1999; Jonides & Nee, 2006; Jonides, Smith, Marshuetz, Koeppe, & Reuter-Lorenz, 1998; Öztekin et al., 2012; Öztekin & McElree, 2007, 2010). Given its well established interactions with memory processes, it is quite conceivable that study material with emotional content might impact forgetting due to PI. For instance, emotion could directly modulate the buildup of interference in memory, resulting in lower levels of interference to resolve. Alternatively, emotion could affect the memory operations that aid the successful resolution of interference.

WM is a dynamic memory system that allows us to maintain a limited amount of information available for processing while processing upcoming information. WM can be described as the cognitive system that acts as a mediator of the encountered information for its short-term maintenance as well as its manipulation to be used while executing a task. Because maintaining information in WM might be an important factor that contributes to the retention of information in the long term, it is expected that emotional information might be treated differently in WM as well.

Studies examining the differences between emotional and neutral stimuli in WM tasks, however, have not capitulated enhancing effects emotion has in long-term memory (Kensinger & Corkin, 2003). For instance, Mather et al. (2006) showed that the source memory (i.e., remembering the relative locations of items) for highly arousing images were poorer than low arousing images, indicating that the arousal component of the stimuli disrupted the inter item binding in WM.

More recently, the impact of emotion on WM has been investigated by presenting emotional stimuli as a task-irrelevant distractor to induce cognitive interference (Denkova et al., 2010; F. Dolcos & McCarthy, 2006). Because emotional information benefits from prioritized processing and tends to capture our attention in a fast and automatic way, presenting it as a distractor might lead to impairments in task-relevant performance (F. Dolcos & Denkova, 2014). When emotional distraction was presented during the delay between to-be-remembered items and test probes in a WM task, specific patterns of brain activity were observed in response to the emotional distractors (F. Dolcos, Diaz-Granados, Wang, & McCarthy, 2008; F. Dolcos & McCarthy, 2006). Impaired WM performance in the presence of emotional distractors was associated with enhanced activity in hot emotional systems, such as the amygdala, and reduced activity in cold executive systems, such as the dorsolateral prefrontal cortex. Research has further shown that when emotional content is task-irrelevant, it can have detrimental effects on memory performance (F. Dolcos & Denkova, 2014; F. Dolcos, Iordan, & Dolcos, 2011; Iordan, Dolcos, & Dolcos, 2013).

Interference of task-irrelevant material impedes our ongoing cognitive functioning. However, forgetting can be overcome if PI is successfully resolved. Accordingly, how we handle PI with emotional material is important in our understanding of the interactions between emotion and memory processes. A common way of inducing interference in experimental settings is using the recent probes (RP) task (Jonides & Nee, 2006; Monsell, 1978). The RP task is an item recognition memory paradigm, where the task demands are the same as a standard item recognition tasks for target test probes. PI is manipulated by changing the recency of the lures. Specifically, presenting a lure from the previous study list induces PI by putting familiarity and source information in conflict. Namely, the recently encountered probe has high residual familiarity and thus is a strong competitor among target representations. To correctly resolve PI, one needs to retrieve detailed episodic information (e.g., list-specific or source information that the item belonged to the preceding trial) in order to correctly reject the recognition probe. Noninterference trials, on the other hand, test a new probe or a probe that was not recently presented, which have low residual familiarity and thus are easy to reject. The detrimental effects of PI on memory performance have been measured by comparing the reaction times (RTs) and false-alarm (FA) rates on interference and noninterference trials (D'Esposito et al., 1999; Jonides & Nee, 2006; Jonides et al., 1998; Öztekin et al., 2012; Öztekin & McElree, 2010). Typically, due to its high residual familiarity, participants are more likely to mistakenly endorse this item, leading to high FA rates on interference trials compared with noninterference trials. Alternatively, participants could successfully resolve PI by engaging in controlled operations that access diagnostic information from memory (e.g., source memory), which leads to slower RTs for interference compared with noninterference trials (Jonides & Nee, 2006; Öztekin et al., 2012; Öztekin & McElree, 2007, 2010).

Levens and Phelps (2008) employed the RP task to investigate how emotion affects memory performance in the presence of PI. Their results indicated lower RT differences between interference and noninterference trials for emotional (E) trials compared with neutral (N) trials, indicating that emotionally arousing stimuli were less affected by PI. Accordingly, they suggested that highly arousing emotional content might act as a facilitator for interference resolution in memory (Levens & Phelps, 2008). There may be two mechanisms underlying this effect: One possibility is that emotion could impact directly on the buildup of interference in memory, by altering the early/automatic familiarity-based responses, and thus leading to lower levels interference to resolve. Alternatively, emotion could impact on the resolution of PI, via modulating the controlled processes that access detailed episodic information from memory. More specifically, there might have been a decrease in the familiarity signal, which would be reflected in the amount of interference created, or an enhancement for the source memory of emotional items that would facilitate the resolution of interference. Unfortunately, RT measures do not allow disentangling the effects on the buildup from resolution of PI. The present study aimed to independently assess the two hypotheses by providing a time-course investigation of how emotion impacts on the buildup and resolution of PI. To do so, we employed a response signal speed-accuracy trade-off (SAT) procedure to the RP paradigm consisting of neutral and highly arousing negative images from the International Affective Picture System (Lang, Bradley, & Cuthbert, 2005), which allowed us to trace the impact of emotion on memory processes that are operative across the full time-course of retrieval (see Figure 2.1). This allowed us to keep track of the impact of emotion in the presence and in the absence of PI, which enabled us to evaluate the use of emotional stimuli when it is task-relevant and task-irrelevant.

SAT Procedure

SAT is a variation of a deadline method in which subjects are signaled to respond at variable intervals after the onset of each test item, allowing a time-course function that measures the growth of retrieval as a function of processing time (McElree & Dosher, 1989). An important advantage of the SAT over traditional paradigms is that it provides conjoint measures of the accuracy and speed of processing, independent of each other. This is in contrast to RT measures derived from traditional tasks, which cannot provide pure measures of processing speed because they are subject to speed-accuracy trade-offs (McElree, 2006). Sampling the full time-course of retrieval also allows independently probing automatic versus controlled operations, because the output of automatic operations have typically been observed to be available before the output of controlled operations across a wide range of tasks (Hintzman & Curran, 1994; McElree et al., 1999; Öztekin & McElree, 2007, 2010; Yonelinas, 2002). Accordingly, the SAT procedure enables independent estimation of both the timing and the magnitude of the output of these processes via quantitative modeling routines (see Figure 2.2 for illustration and description of a hypothetical SAT function). Previous studies investigating nonemotional PI with SAT (Hintzman & Curran, 1994; McElree & Dosher, 1989; Öztekin et al., 2012; Öztekin & McElree, 2007, 2010) have shown that recent negative (RN) probes induce high FA rates early in retrieval compared with nonrecent or unstudied lures. However, crucially, the elevated FA rates diminish later in retrieval when participants are able to recover more detailed episodic information (i.e., either that the probe was not a member of current study list or that it was studied on previous trial). This nonmonotonic FA function for RNs is consistent with two operations in opposition: Automatic assessments of familiarity engender high FA rates early in retrieval, which are then subsequently countered by controlled, strategic retrieval operations that serve to recover detailed episodic information.

Figure 2.3 illustrates a hypothetical two-phase retrieval function that characterizes the buildup and resolution of PI. Buildup of PI starts at the point in time when item familiarity becomes available (see Point A in Figure 2.3). During this time, performance shows an exponential increase in FA rates reaching an asymptote until the time when diagnostic information (e.g., source or list-specific information) becomes available. When given enough time, participants can engage in controlled retrieval operations and access the relevant source information (i.e., that the RN probe was not in the current study list). After this breakpoint (see Point B in Figure 2.3), PI can be resolved and FA rates diminish. Accordingly, evaluating how the buildup and resolution of PI unfolds across the full time-course of retrieval allowed us to assess whether emotional content of the material directly affects the early/automatic period where interference builds up, or the controlled memory operations that successfully resolve PI. As such, this fine-grain analysis has the potential to identify the underlying mechanisms that modulate the relationship between emotion and forgetting in memory.

2.3. Method

Participants

Twenty students (14 undergraduates and 6 graduate students) from Koç University participated in the experiment. All participants gave written informed consent before participation. Two participants were members of the lab and volunteered their time. The remaining participants were compensated for their time. Data from four participants who failed to comply with the SAT procedure were excluded from analyses, leaving 16 participants. SAT procedure employs within-subject comparisons across experimental conditions with intensive data collection from each participant (e.g., the present study consisted of over 2,000 experimental trials per participant). Accordingly, consistent with previous investigations (McElree & Dosher, 1989; Öztekin et al., 2012; Öztekin & McElree, 2007, 2010), the present study aimed for a conservative sample size of 16 participants.

Design and Stimuli

The experiment consisted of twelve 20-min sessions, completed over several weeks. Each session contained 84 emotional and 84 neutral experimental trials in which participants studied a three item list consisting of images. Next, participants were cued to respond to a recognition probe following a brief visual mask and indicated whether the test probe was a member of the study list using a “yes” or “no” recognition response. Participants completed a 20-min practice session with nonemotional stimuli to train for the SAT procedure.

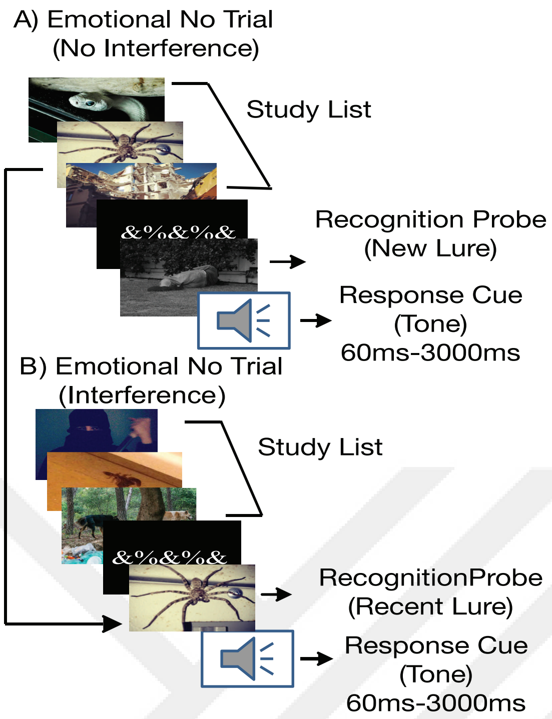

Figure 2.1. Illustration of non-interference (Panel A) and interference (Panel B) trials for emotional stimuli. For each trial, 3 study images were presented, followed by a visual mask. Participants indicated whether the recognition probe was presented in the study list or not. They were trained to respond within 300ms of the response signal.

Study design adapted a response-deadline version (McElree & Dosher, 1989; Öztekin et al., 2012; Öztekin & McElree, 2010) of the RP task, a widely used paradigm to induce PI by presenting lures from previous study list. The stimuli were chosen from the International Affective Picture System (Lang et al., 2005), according to the standardized scores of valence and arousal levels. For the emotional stimuli set, 273 highly arousing (*M* = 5.80, *SD* = 0.69) and negatively valenced (*M* = 2.85, *SD* = 0.69) images were selected. The neutral stimuli set consisted of 273 low arousing stimuli (*M* = 3.5, *SD* = 0.63), with medium valence levels (*M* = 5.26, *SD* = 0.76). The emotional stimuli set was significantly more arousing, *t*(272) = 40.673, *p* < .001, and more negatively valenced, *t*(272) = 38.792, *p* < .001, than the neutral stimuli set.

Each image was presented once per session. There were equal numbers of targets; probes that required a “Yes” response and lures; probes that required a “No” response for both (E) and (N) trials. Targets were chosen equally from the three serial positions (SPs) of the current study list. Half of the lures consisted of recently presented lures selected from the study list of previous trial (recent lures), and half consisted of lures that had not been presented in the current session (new lures).

Procedure

Figure 2.1 illustrates the sequence of events in a single trial: (a) Study images were presented sequentially for 1200 ms each. (b) The study list was followed by a visual mask, consisting of nonletter symbols for 1200 ms. (c) After the mask, the test image was presented for the duration of the response deadline. (d) At 60, 200, 300, 500, 800, 1500, or 3000 ms after the onset of the recognition probe, a 50-ms tone sounded to cue participants to respond. (e) Participants indicated provided a yes-no recognition response as quickly as possible after the onset of the tone by pressing a key. (f) After providing a response, participants were given feedback on their latency to respond. Participants were trained to respond within 300 ms of the tone. They were informed that responses longer than 300 ms were too slow and responses less than 100 ms were anticipations, and that both should be avoided. (g) After the latency feedback, participants were asked to give a confidence rating ranging from 1 (low confidence) to 3 (high confidence). The confidence ratings primarily served to enable participants to self-pace themselves through trials and were not analyzed. Participants initiated the next trial by pressing a key.

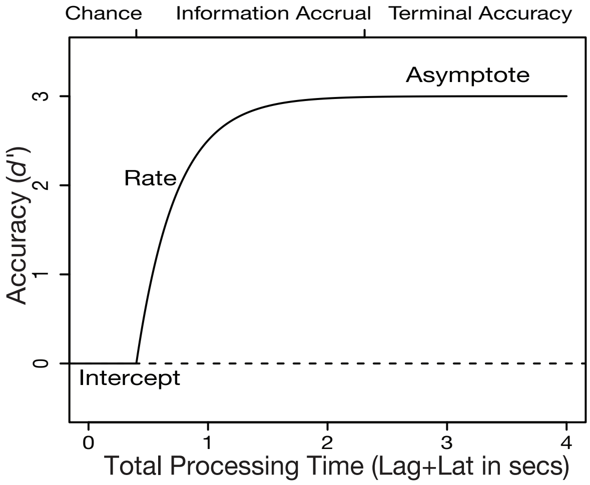

Figure 2.2. Illustration of a hypothetical SAT function that shows how accuracy (in d' units) grows over processing time (in seconds). The SAT curve reflects three phases: A period where performance is at chance (the departing point in time from chance is marked by the intercept parameter), followed by a period of information accrual (the rise of this information accumulation is reflected by the rate parameter of the SAT function), and following this period, the maximum level of accuracy is reached, where performance does not improve any more (the asymptote parameter of the SAT function).

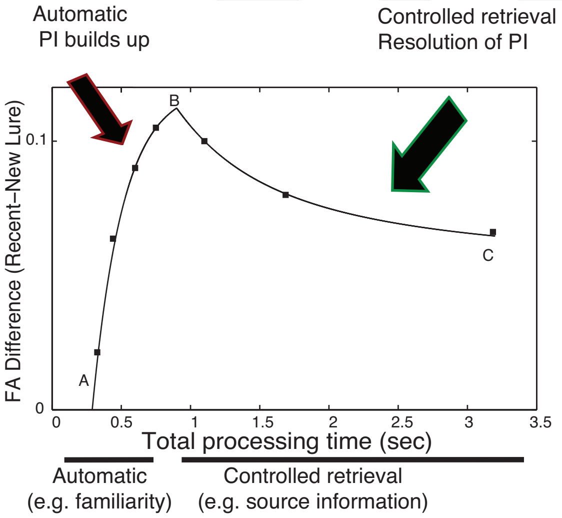

Figure 2.3. Illustration of a hypothetical two phase retrieval function that shows the behavior of false alarm difference between interference and non-interference trials. The function reflects; 1) the time point when participants start to make FAs (familiarity intercept; point A), 2) the speed of interference build-up (rate parameter), 3) the terminal level FA difference reaches before engaging in controlled processes and correct their judgments (familiarity asymptote; point B), 4) the time point when participants start to recover the detailed/contextual information and correct their judgments (recollective intercept; point B), 5) the terminal level of FAs after the retrieval of source information (recollective asymptote; point C).

2.4. Results

For targets, each participant's hit rates were scaled against the FA rates to new lures (NL) to obtain (equal-variance Gaussian) *d′* measures. To ensure measurable *d′*s, we adapted a minimal correction procedure as suggested by Snodgrass and Corwin (1988).

General effects of Emotion on the Dynamics of Memory Retrieval; in the absence of Interference

Asymptotic Accuracy

Consistent with previous research, we averaged *d′* for the last two response deadlines in order to obtain empirical measures of asymptotic accuracy (Fig. 2.4). This measure reflects the terminal accuracy level reached, indicating the probability of successful retrieval (McElree, 2001; McElree & Dosher, 1989, 1993; Öztekin & McElree, 2007). A 2 (stimuli type [Emotion vs. Neutral]) x 3 (serial position of positive test probes) Analysis of Variance (ANOVA), conducted on asymptotic *d′* s indicated a main effect of serial position (SP) for both Emotion (E) and Neutral (N) trials: asymptotic accuracy increased as a function of SP of the test probe [E; $F(2,15) = 12.\ 037$, $p < .005$, $\eta_p^2 = 0.445$ N; $F(2,15) = 9.177$, $p < .005$, $\eta_p^2 = .380$]. Post hoc comparisons indicated that the mean score for the asymptotic accuracy of later serial positions were significantly higher for both emotional (SP1; M=3.58, SD=0.43, SP2; M= 3.81, SD=0.21, SP3; M=3.96, SD=0.17) and neutral (SP1; M=3.66, SD=0.4, SP2; M= 3.85, SD=0.19 SP3; M=4.02, SD=0.14) material. On the other hand, there was no measurable impact of emotion on asymptotic accuracy ($p > .315$, $\eta_p^2 = .067$), participants performed comparable across emotion and neutral trials. We also did not observe a reliable stimuli type by SP interaction ($p > .915$, $\eta_p^2 = .006$).

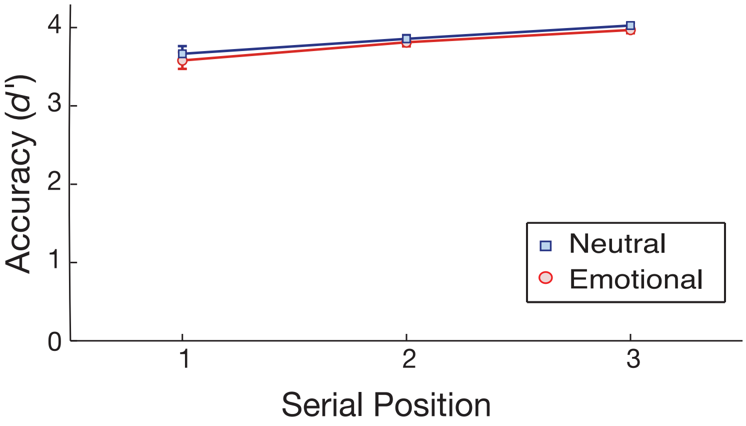

Figure 2.4. Asymptotic accuracy levels (averaged d' over the last two response deadlines) across the three study positions. Error bars denote standard error of the mean (SEM).

Retrieval Dynamics

We estimated the retrieval dynamics by fitting the individual participants' data and the average data (derived by averaging d' values for each condition across participants) with an exponential approach to a limit:

$$d'(t) = \lambda\left(1 - e^{-\beta(t-\delta)}\right), t > \delta, \text{else } 0. \tag{2.1}$$

In Equation (2.1), *d'(t)* is the predicted d' at time *t*; λ is the asymptotic accuracy level reflecting the overall probability of recognition; δ is the intercept reflecting the discrete point in time when accuracy departs from chance ($d' = 0$); β is the rate parameter, which indexes the speed at which accuracy grows from chance to asymptote. Previous studies have indicated that this equation provides a good quantitative summary of the shape of the SAT functions (Dosher, 1981; McElree & Dosher, 1989; Wickelgren & Corbett, 1977; Wickelgren et al., 1980).

The quality of the model fits were assessed by: (a) The value of an adjusted-R^2 statistic (Reed, 1973); (b) the consistency of parameter estimates across participants;

(c) evaluation of whether the fit yielded systematic deviations that could be accounted for by additional parameters. These latter two metrics were assessed by statistical tests conducted on the parameter estimates derived from the model fits across participants.

Within and Composite List Dynamics

We first evaluated whether SAT functions for neutral and emotional trials exhibited the same patterns observed in previous studies (McElree, 2006). To do so, SAT functions for the 3 SPs were fit with sets of nested models that systematically varied the 3 parameters in Equation 2.1. Our results indicated that the most recent item benefits from a privileged state in the focus of attention (Mcelree & Dosher, 1993; Öztekin et al., 2012; Öztekin & McElree, 2007, 2010; Wickelgren et al., 1980) and extends this phenomenon to emotional study material (Supplemental Material 2 for these results in more detail).

We next evaluated overall differences in terminal accuracy and retrieval speed across neutral and emotional stimuli. To do so, we averaged individual participants' as well as the average *d'* values across the three serial positions. These composite list functions were then fit with Equation 2.1 as described above. Fig. 2.5 illustrates the SAT functions for the average E and N data, with smooth curves indicating the fitted exponential functions.

Consistent with the pattern observed in within list dynamics for emotional and neutral trials, analyses of the retrieval dynamics across E and N composite SAT functions revealed a reliable difference between E and N in retrieval speed measures. Specifically the intercept parameter, which marks the point in time when performance departs from chance, was slower for emotional compared to neutral trials [$t(15) = -2.50$, $p < .025$, $d = 0.631$]. The average (across participants) intercept parameter was

296 ms and 308 ms for N and E respectively (Tables S2.3A and S2.3B present the parameter estimates for individual participants and the average data).

In contrast to retrieval dynamics measures, terminal accuracy levels, as assessed by the asymptote parameter were comparable across emotional and neutral trials ($p > .85$, $d = 0.051$). Thus, the impact of emotion was prominent on retrieval speed measures, with no measurable impact on probability of successful retrieval for positive trials, in the absence of interference.

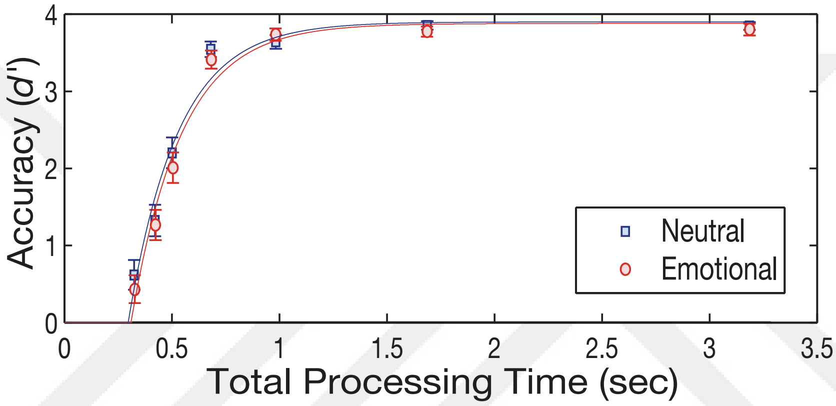

Figure 2.5. Illustration of composite list SAT functions and model fits. Accuracy (in *d′* units) for composite list (averaged over serial position of the test probe) SAT functions plotted against total processing time (duration of the response deadline plus latency in seconds) for the average (over participants) emotional and neutral trials. The symbols indicate empirical data points, and the smooth lines indicate the model fits derived from Eq. (2.1). Error bars denote SEM.

Effects of Emotion on the Dynamics of Memory Retrieval; in the presence of Interference

FA Analyses: Effects of Emotion on the Build Up and Resolution of Interference

We now turn to our analyses of the false alarms to recent and distant lures to see whether and how participants’ response patterns to reject lures differ across E and N trials. A comparison of the response patterns across the whole time-course allowed us to independently investigate the differential impact of emotion on the build up and resolution of PI. To do so, we analyzed the differences in false alarm rates between our recent negatives (RNs), lures that were studied in the preceding study list, and distant negatives (DNs), lures that had not been studied in the current session. This measure provided an unbiased measure of performance by factoring out participants' bias to endorse a test item as member of the study list (e.g. a tendency to respond yes more often than no, regardless of the type of probe).

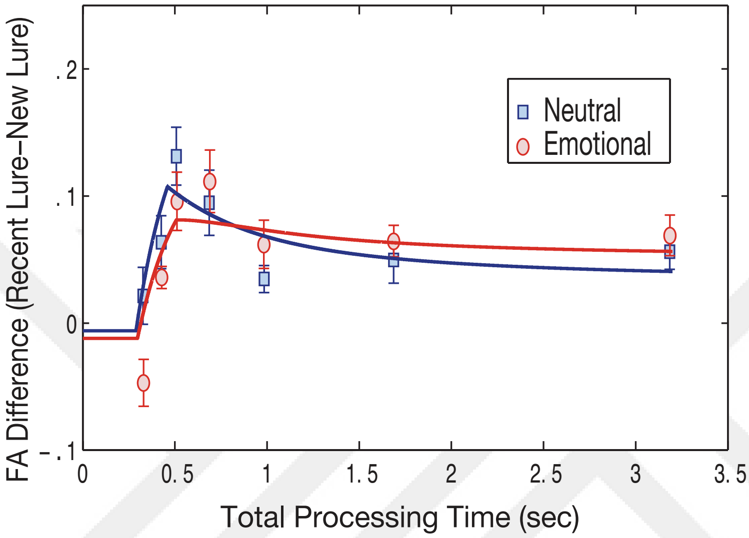

Figure 2.6. Difference in false alarm rates across the recent and distant negative probes plotted against total processing time (duration of the response deadline plus latency in seconds) for the average (across participants) of emotional and neutral trials. The symbols indicate empirical data points, and the smooth lines indicate the model fit derived from Eq. (2.2). Error bars denote SEM.

Note. RN = recent negative; DN = distant negative

Fig. 2.6 plots the FA difference scores for the average E and N trials. Note that due to the FA difference scaling, higher scores indicate a higher tendency to false alarm to RNs. Fig. 2.6 shows that for both emotional and neutral stimuli, the FA difference scores increase early in retrieval and then diminish later in retrieval. This non-monotonic pattern indicates that the information basis for the recognition memory judgments has shifted across retrieval and is consistent with previous research (Öztekin et al., 2012; Öztekin & McElree, 2010). The early high FA rates presumably indicate the contribution of familiarity (because the RN has been studied on the previous trial, it has high residual familiarity compared to the DN). The observed reduction in FA rates later in retrieval suggests the accrual of new information that contributes to the recognition judgments, presumably reflecting source or list specific information recovered by a recollective process (e.g., the fact that the RN probe was studied on previous trial, or that it was not a member of the current study list). Below, we statistically assess these differences with a modified version of the quantitative two-process SAT model originally suggested by Ratcliff (Ratcliff, 1980) and adapted to the exponential form by McElree & Dosher (1989).

$$FA_{diff}\ (t) = \begin{array}{ll} [\lambda\ (1 - e^{-\beta(t-\delta_1)})] + \gamma, & \text{for } \delta_1 < t < \delta_2 \\ [\lambda_2 + (\lambda_1 - \lambda_2)(\delta_2 - \delta_1)/(t - \delta_1) * (1 - e^{-\beta(t-\delta_1)})] + \gamma, & \text{for } t \geq \delta_2. \end{array} \tag{2.2}$$

Equation (2.2) states that during the initial retrieval period ($\delta_1 < t < \delta_2$), accuracy depends on accrual of one type of information, presumably familiarity information. During this initial period, accuracy is modeled by the top portion of Equation (2.2), a simple exponential approach to an asymptote (λ_1). At time δ_2, a second source of information starts to contribute to the recognition memory judgments. This source of information could arise from the output from a second process, e.g., a recollective operation that accesses detailed episodic information. The accrual of this second type of information leads to the change in retrieval; shifting the asymptote from λ_1 to λ_2.The bottom portion of Equation 2.2 states that response accuracy gradually shifts to the new asymptote (λ_2) starting at time δ_2. In addition, a shifting parameter (γ) was added to the model to account for negative scores[4].

We should state that the two processes mentioned in the aforementioned model presented in Eq. (2.2) contemplate the recovery of information through an automatic versus controlled retrieval operation. These processes need not be equivalent to familiarity versus recollection. While the notion of a fast/automatic familiarity assessment (or a component of stimulus/item identification that could lead to an increase in false alarm rates) is consistent with dual process theories of recognition (e.g., see Yonelinas, 2002 for review), the slower controlled component might not necessarily equal general recollection, but could also reflect the independent accrual of diagnostic episodic information (e.g. source information) that can aid the successful resolution of interference. The slower accrual of this diagnostic episodic information can overrule the contribution of the fast/automatic assessments (independent of whether the automatic process has reached completion), leading to the nonmonotonic pattern observed in the data.

To test the effects of emotion on the familiarity-based responses that reflect the build up of PI versus the controlled processes that reflect the resolution of PI, we fit each participants' and average E and N data with Eq. (2.2) and compared the parameter estimates derived from each phase. Namely, if E directly impacts the build up of PI, we should see differences early in retrieval, either on the magnitude or timing of familiarity-based judgments. If on the other hand E impacts on the controlled operations that aid successful resolution of PI, then we should observe differences later in retrieval. These effects can be observed either on the magnitude or

---

[4] Due to the observed negative scores in the emotional data, future research would provide additional insight as to whether a revised two-process retrieval model would be more appropriate for emotional material.

timing of responses based on controlled, responses that access detailed episodic information and aid resolution of PI.

A paired *t*-test comparison across the parameter estimates for emotional and neutral trials indicated that the familiarity asymptote (λ_1; 0.154 for average E, and 0.22 for average N) was marginally lower ($p < .091$, $d = 0.470$) for emotional compared to neutral material. Crucially, the recollective asymptote (λ_2; 0.075 for average E, and 0.032 for average N) was marginally higher ($p < .056$, $d = 0.69$) for emotional compared to neutral material. Notably, this crossover (Asymptote [λ_1 vs λ_2] by Condition [Emotional vs Neutral]) interaction was significant [F = 4.71, p < .046]. Additionally, there was a speed difference for the information accrual, reflected by the rate parameter estimate, t (15) = 2.77, $p < .014$, $d = 0.918$ (163 ms in $1/\beta$ units for average E and 116 ms in $1/\beta$ units for average N), indicating that the buildup of interference was *slower* for emotional study material. The two intercept estimates, however, (δ_1 and δ_2) were comparable across emotional and neutral stimuli (δ_1; E = 330 ms, N = 344 ms, $p > .17$, $d = 0.11$; δ_2; E = 650 ms, N = 494 ms, $p > .094$, $d = 0.62$). Tables S2.4A and S2.4B present the parameter estimates for individual participants' and average data for E and N.

These data suggest that early in retrieval when processing is largely automatic, emotion has a beneficiary effect on memory performance by slowing the build up of interference. However, in contrast, our data indicate that despite leading to a slower build up of PI, emotion leads to less effective resolution of PI later in retrieval when retrieval is more based on controlled retrieval operations that access diagnostic information from memory. Thus, the data point to the conclusion that emotion differentially impacts the build up and the resolution of interference in memory.

2.5. Discussion

The present study aimed to provide an in-depth investigation of the relationship between emotion and forgetting. To do so, we tracked the impact of emotion on memory performance in the absence and presence of interference across the full time course of retrieval, allowing us to independently examine its impact on the buildup and resolution of interference, as well as its general effects on the dynamics of memory retrieval in the absence of interference. Prior work on the effects of emotion on interference resolution has suggested that there might be a facilitatory effect of having an emotional component (Levens & Phelps, 2008, 2010). In the current investigation, we endeavored to uncover the exact mechanisms behind this effect emotion has on PI and its resolution in memory in order to provide new insights on how emotion impacts forgetting.

General Effects of Emotion on Retrieval Dynamics in the Absence of Interference: Distinct Impact on Retrieval Success and Retrieval Speed

In the absence of interference, both the composite list and the SP analyses indicated comparable asymptotic accuracy levels for neutral and emotional study material. However, crucially, retrieval dynamics estimates indicated a slower intercept, that is, the point in time when information first becomes available, for emotional compared with neutral study material. Recognition memory judgments are thought to be based on the accrual of two types of information: a fast assessment on the quality of the match of a probe to representations in memory (often referred to as familiarity) and a rather detailed contextual information recovery of information (such as source information). Recovery of source information, often viewed as recollection, is usually led by effortful retrieval operations and, as a consequence, accumulate slower than familiarity based assessments that are thought to be fast and more automatic (Yonelinas, 2002). Although it is not possible to separate the contributions of familiarity and source information for positive trials, previous time-course investigations have shown that familiarity information becomes available earlier than detailed episodic information (McElree et al., 1999; Öztekin & McElree, 2007). Hence, the intercept difference between emotional and neutral stimuli observed in our study presumably reflects the effect of emotional content on fast/automatic familiarity assessments.

Relationship Between Emotion and Forgetting: Differential Impact of Emotion on the Buildup and Resolution of Interference in Memory

Manipulating the presence of interference in the retrieval context allowed us to investigate the relationship between emotion and forgetting. Specifically, comparison of responses to RN probes, lures studied on the previous study list, and hence with high residual familiarity, with DNs, lures that had not been studied in an experimental session, allowed us to track the differential impact of emotion on the time-course of responses early and later in retrieval; that is, responses based on familiarity that are dominant early in retrieval and reflect the buildup PI, and those based on recovery of detailed information that are dominant later in retrieval and aid in successful resolution of interference in memory.

Emotional information is treated differently due to its survival value, hence, the importance of how a task-irrelevant emotional memory representation could interfere with recognition memory judgments. Our FA analysis showed that: (a) the buildup of PI for emotional material was slower (indicated by a slower rate parameter in the modeling estimates) compared with neutral material; (b) emotional material did not lead to as many FA difference as neutral material did (reflected as a lower familiarity asymptote); and (c) however later in retrieval, the resolution of PI was less effective for emotional (observed as a higher recollective asymptote) than neutral material.

Our findings indicating a slower buildup and a lower amount of PI for emotional material are consistent with the delayed onset of information accrual for positive trials. Taken together, the data indicate that the impact of emotion on memory processing is largely dominant early in retrieval during automatic processing, presumably reflecting the contribution of familiarity-based judgments. On the other hand, we also observed an effect of emotion on the controlled memory operations that aid successful resolution of PI. However, this effect was not in favor of emotional stimuli. Conclusively, our FA analysis revealed that while once goal relevant but yet irrelevant emotional stimuli reduce the effects of PI early in retrieval, its resolution is not as effective later in retrieval. We further discuss these findings below.

Effects of emotion on automatic processes early in retrieval (buildup of PI). Our time-course investigation revealed that the previously observed (Levens & Phelps, 2008) facilitatory effect emotion has on the control of PI emerges early in retrieval as a slower rate of accrual for distracting material, leading to a reduced amount of PI to resolve. Previous work, which investigated the neural circuitry behind this facilitatory effect, has shown that although resolution of interference is associated with enhanced activation in the inferior frontal gyrus for both emotional and neutral items, interference resolution of emotional items recruited additional regions such as anterior insula and orbital frontal cortex (Levens & Phelps, 2010). Most importantly, the left amygdala showed differential activity for emotional stimuli. Previous work has also shown that effects of emotion have relatively earlier traces in special limbic areas such as the amygdala, acting as a gating mechanism for sensory processing, projecting to distant regions at later latencies (see Pourtois, Schettino, & Vuilleumier, 2013 for a review). As such, the amygdala is seen as a mediator for emotional effects along with its interactions with frontal and temporal brain structures (LaBar & Cabeza, 2006). Consistent with an early effect of emotion led by the amygdala, Düzel and colleagues (Fenker, Schott, Richardson-Klavehn, Heinze, & Duzel, 2005) observed right amygdala activation for Know (subjective basis of judgments related to familiarity) but not for Remember (subjective basis of judgments related to recollection of specific details) responses during recognition of emotional faces. Consistent with prior work that has indicated the impact of the amygdala to be dominant early in processing, it is possible that, in the current investigation, amygdala activity might have delayed the accrual of familiarity information, the dominant source of information early in retrieval.

Additionally, previous research (e.g., Denkova et al., 2010; Dolcos et al., 2008; Dolcos & McCarthy, 2006) has indicated a distinct pattern of activity in the systems specific to emotion processing and to executive processes. Specifically, when emotional distraction was presented during the delay between to-be-remembered items and probes in a WM task, an increased pattern of activity was observed in the amygdala, while activity in the dorsolateral prefrontal cortex decreased for emotional material. The differential activity pattern in the emotion-related and executive systems might also be prominent in the control of PI of emotional material. Studies investigating the effects of emotional interference on cognitive functioning have indicated that task performance depends on the dynamic interactions between neural networks that supports emotion processing and the brain regions that assist the efficient use of task-relevant information (F. Dolcos, Denkova, & Dolcos, 2012). One might speculate that heightened amygdala activity early in retrieval interferes with the automatic processing of item information (i.e., lowered or delayed perirhinal cortex activity) resulting in reduced availability, hence, less interference. This pattern, however, might be specific to the projections early in retrieval, which might change later in retrieval. Consistent with this notion, several studies (Pitkänen, Pikkarainen, Nurminen, & Ylinen, 2009; von Bohlen und Halbach & Albrecht, 2002) have provided evidence for strong projections between the amygdala and the perirhinal cortex. Future neuroimaging studies focusing on early amygdala and hippocampal structure interactions could provide further insight into the effects of emotion on memory processing early in retrieval. Effects of emotion on the controlled processes later in retrieval (resolution of PI). When resolving the conflict in the presence of PI, one needs to rely on the relevant information in the competing memory representations, namely, the familiarity of the item encountered with an item from the preceding trial, and more specific episodic information that determines the item belongs to the previous study list. The successful resolution of PI requires accessing and selecting the diagnostic source information that the probe belongs to the previous trial and discounting the familiarity information. During this phase, our data showed that, for emotional study material, PI was not resolved as successfully as neutral material. That is, the asymptote reached after the resolution of PI has been completed was higher for emotional items, presumably indicating a reduced availability of recollective/detailed episodic information.

Although this finding might seem at odds with the facilitative effects of emotion on PI noted earlier (e.g., Levens & Phelps, 2008, 2010), it is important to note that, in traditional RT experiments where processing time is not controlled, it is not possible to capture the entire retrieval process: Most typically, participants will respond fast, thus allowing capture mostly of the buildup of PI, but not allowing capture of its complete resolution. However, our time-course investigation provided the retrieval pattern in which the two phases of PI can be fully observed. We next discuss possible explanations for the distinct pattern observed during the resolution of PI.

The RP task is a manipulation of an item recognition task in which participants are asked to indicate whether a probe was presented in the current study list. Participants are not informed about the recency manipulation, hence, the sole aim of the task for them is to correctly recognize the encounter of the item in a trial. Accordingly, contextual information about the item (i.e., the item belonging to the trial n or trial n-1) is not central to the task requirements. In the case of emotional material, because the contextual information is not the main focus of the task, there might be a trade-off between central and peripheral details of the item. Previous work has suggested that such a trade-off (i.e., emotion induced memory trade-off) arises due to the fact that processing of emotion narrows attention, and that the mnemonic signal for the emotional components of the item is strengthened, leaving nonemotional details unattended (Murray & Kensinger, 2012; Phelps & Sharot, 2008). However, although strengthened emotional details can have a boosting effect on the subjective sense of remembering, it does not guarantee a general enhancement of recollection (Rimmele, Davachi, Petrov, Dougal, & Phelps, 2011; Sharot & Yonelinas, 2008). In general, vivid and enhanced recollective experience is accompanied by successful recovery of contextual details of the event, which reflects the recollection component of recognition (Yonelinas, 2002). It is important to state that although emotion is expected to have an enhancing effect on memory in general, in this case, there appears to be a discrepancy between the enhanced subjective sense of recollection and memory for contextual details of emotional material.

When emotion’s effect on contextual and relational information are examined, a specific pattern emerges: emotion enhances source memory for the features that are perceptually bound to the item while it impairs contextual/relational details that are not inherent to the item (see Chiu, Dolcos, Gonsalves, & Cohen, 2013 on opposing effects of emotion on contextual or relational memory; see Kensinger, 2009 for a review on emotion's effects on remembering the details). For instance, it has been shown that the font color or spatial location information of emotional items are remembered better than neutral items (D'Argembeau & Van der Linden, 2005; Doerksen & Shimamura, 2001; Mather & Nesmith, 2008). On the other hand, a group of studies has found impaired detailed memory of a scene/context in which emotional items have been embedded and presented (see Kensinger, Garoff-Eaton, & Schacter, 2006a).

Additionally, other investigations have shown that relational binding of the emotional item pairs is worse than neutral items (Knight & Mather, 2009; Nashiro & Mather, 2011; Pierce & Kensinger, 2011). Findings from a recent study conducted by Bisby and Burgess (2014) supported this differential impact of emotion on item and associative memory; contextual memory of emotional items were impaired compared with neutral items while memory for individual emotional items were enhanced or unaffected under certain conditions. Conclusively, it has been suggested that the emotional element of an object induces enhanced focused attention that alters within-object binding (i.e., features related to the object such as color or location) while this narrowing of attention might not favor the object-object bindings or the binding of contextual features (Mather, 2007). The pattern we observed in our study is consistent with this recent work, and indicates that there might be a cost to having an emotionally arousing component that reduces availability of nonprioritized episodic information, such as source or contextual memory, which, in our case, was list-specific episodic information required to successfully resolve PI.

2.6. Conclusion

The present study provides an in-depth investigation of the relationship between emotion and forgetting via tracking the full time-course of how the impact of emotion on memory unfolds in the presence of interference. Our findings suggest that emotion has a differential impact on the buildup and resolution of PI in memory: Emotion aids memory by slowing down the buildup of interference and leads to less interference early in retrieval when memory judgments are largely automatic. This finding provides an explanation for the previously observed facilitating effect of emotion during PI. However, this facilitating effect should be interpreted cautiously. In contrast to the pattern observed during the buildup of PI, our data indicate that emotion has a negative impact later in retrieval when PI is resolved via controlled memory operations that access detailed episodic information from memory.

CHAPTER III

Neural and Behavioral Patterns of Proactive Interference for Emotion and Neutral Information in Working Memory

3.1. Abstract

Proactive interference (PI) is the tendency for information learned earlier to interfere with more recently learned information. In the present study we induced PI by presenting items from the same category over several trials. This results in accumulation of PI and reduces the discriminability of the items in each subsequent trial. We introduced emotional (e.g., disgust) and neutral (e.g., furniture) categories and examined how increasing levels of PI affected performance for both stimulus types. Participants were scanned using functional magnetic resonance imaging (fMRI) performing a 5-item probe recognition task. We modeled responses and corresponding response times with a hierarchical diffusion model and showed that PI effects on latent processes (i.e., reduced drift rate) were similar for both stimulus types. But the effect of PI on drift rate was less pronounced PI for emotional compared to neutral stimuli. For neutral stimuli, but not for emotional stimuli, the decline in drift rate was accompanied by an increase in neural activation in parahippocampal regions. Recovery from PI was mediated by anterior VLPFC activation for emotion trials.

3.2. Introduction

One common scenario that makes forgetting noticeably annoying in our daily lives is when we set up a new password and then use the old version the next time we sign into our account. Only after failing once or twice, we remember that we updated our password and retrieve the new version. We often suffer from retrieval failures as such due to interference of previously learned yet irrelevant information (i.e., the old password), which is known as proactive interference (PI). PI is a huge cost on our cognitive system, so much that it has been suggested forgetting from working memory would be minimum if there was no interference (Jonides & Nee, 2006).

Although retrieval failures seem like a common occurrence, we are surprisingly good at dealing with PI, ergo, understanding mechanisms behind the control of PI is important. A vast amount of research evaluated how we efficiently retrieve relevant information from our working memory in the presence of PI (Badre & Wagner, 2005; D'Esposito et al., 1999; Jonides et al., 1998; Öztekin & Badre, 2011; Öztekin et al., 2012; Öztekin & McElree, 2007). While these studies provided rich findings suggesting that we adapt retrieval strategies and gear up with controlled retrieval processes to resist PI, these effects were mostly tested and demonstrated with neutral stimuli. Yet a significant amount of time in our lives we either deal with emotionally loaded stimuli or we are subject to physiological changes in our body when we are stressed or anxious. The likelihood of retrieving the new password could be higher or lower depending on its emotional salience or if the outcome of retrieving the correct password was highly important. For example, imagine that you received a notification e-mail saying that your account has been used by an unrecognized source. Prior work on emotion memory interactions showed that emotional events are remembered better and more resistant to forgetting than neutral events (see Buchanan, 2007; Hamann, 2001; LaBar & Cabeza, 2006, for review). Hence it is critically important to understand how emotion alters our retrieval processes that are adapted to deal with PI.

In a recent study, Mizrak and Öztekin (2016) showed that PI from arousing stimuli and PI from neutral stimuli has different effects on working memory retrieval dynamics. They used the Recent-Probes (RP) task which is commonly used to induce PI in experimental settings (Monsell, 1978). In this task, participants are asked to hold a small set of target items in memory over a brief retention interval and then make a *yes* or *no* decision to a recognition probe indicating whether the probe was one of the target items or not. In the RP task, the recognition probe can be an item from the study list (target), an item that participants did not encounter before (lure) or an item that was presented in the previous trial but not the current one (recent lure). Typically, participants are more prone to falsely recognize the recent lure as a target item compared to non-recent lures. If the temporal information of the recent lure (the previous study list) is recovered via controlled retrieval processes, PI can be easily overcome. Mizrak and Öztekin (2016) investigated the full time course of retrieval in which PI emerges and builds up due to early item information and is resolved via retrieving the list membership of the probe later in time. Their findings revealed an interaction of PI time course and stimulus type. The build up of PI for emotional trials was slower than for neutral trials, leading to lower amounts of interference early in retrieval. However, detrimental effects of PI were stronger for emotional compared to neutral trials later in retrieval. There is growing evidence showing that emotion slows down forgetting from long term memory (Yonelinas & Ritchey, 2015) and Mizrak and Öztekin (2016) study extended these findings to working memory.

In the current study, we aim to test the impact of emotion on PI in a different experimental setting. Another way of inducing PI is by presenting many semantically/categorically similar items. In the classic release-from-PI (rfPI) paradigm Wickens (1970) manipulated PI by presenting items from the same semantic category over several trials and by switching categories between the sets of trials. When the items share the same features, they are highly similar to each other which reduces the discriminability of the items (Nairne, 2002). More specifically, a particular cue that can be used in retrieval will now be related to multiple items which leads to a cue overload. As a result, the items to be retrieved in the latter trials of the consecutive presentation are subject to more competition, which in turn leads to performance decline (i.e., slower responses, lower memory accuracy). Studies that used this manipulation consistently showed that; detrimental effects of PI are eliminated when the category is changed and the items presented are dissimilar to the items presented in previous lists (Öztekin & Badre, 2011; Oztekin et al., 2009; Öztekin & McElree, 2007; Watkins & Watkins, 1975; Wickens, 1970). Participants can overcome the PI effects on memory performance by selectively retrieving unique/relevant details about the probe which requires cognitive control.

Here, we manipulated the content of the categories. We introduced emotional (i.e., disgust) and neutral (i.e., furniture) categories and examined how varying PI affected performance when items were drawn from emotional categories compared to neutral categories. For emotion categories, we used images that belong to the same category of emotions (i.e., disgust vs. fear) which are highly arousing and unpleasant. It is important to study how cognitive and neural mechanisms behind the processes involved in resolving PI are altered by emotion. Hence, in addition to behavioral measurements, we also examined the neural responses during emotion and neutral trials given to different levels of PI.

In the aforementioned study which measured PI via inducing episodic similarity (Mızrak & Öztekin, 2016), it has been shown that emotion had a differential impact on the build up of PI as well as the resolution process. While emotion led to lower levels of interference to be resolved by slowing down familiarity assessments, it hindered the controlled processes that were required to retrieve the temporal context of the item. When emotion's effect on contextual and relational information are examined, a specific pattern emerges: emotion enhances source memory for the features that are perceptually bound to the item while it impairs contextual/relational details that are not inherent to the item (see Chiu et al., 2013 for the opposing effects of emotion on contextual memory; and Kensinger, 2009). While contextual details of the item that were not bounded to the item itself were impaired for emotion items, visual details that were central to the item were remembered better compared to neutral memoranda (see Kensinger, Garoff-Eaton, & Schacter, 2006b). Recently, Yonelinas and Ritchey (2015) proposed an account that provides a partial explanation to the observed findings in this literature. They suggested that emotion enhances recollection of item-emotion bindings mediated by amygdala which might be more resistant to forgetting than hippocampus-dependent item-context bindings. Collectively, we can infer that emotion might increase the likelihood of remembering item details rather than context details.

In the present study, we expected participants to use the specific emotion an item elicits (disgust vs fear) as a retrieval cue. Participants might use this shared cue to indicate the item as a target or lure, however in the latter trials, along with the current study list items, previously studied memoranda will also be retrieved under the same cue. Since there are many items that come from the same category as the probe, participants would not benefit from relying on the category membership but retrieving more diagnostic features that are unique to the item. In the case of emotion, we expected activated items to be retrieved with more item-related details than neutral memoranda due to the enhanced item-emotion bindings. Thus, even if participants are inclined to base their decisions on the elicited similarity of the item, these item-specific details can be used to recover from the cue overload. Accordingly, we hypothesized that for emotion trials participants would be subject to PI and memory performance would be affected by competition between similar items but they should overcome PI effects for emotion trials easier/better compared to neutral trials. Such a finding would contradict with our previous findings which showed controlled retrieval process to recover from PI was less efficient for emotional stimuli. However, this is due to different types of information (item vs context) that is required to overcome PI across the previously used recent-probes task and rfPI task.

The ability to correctly discriminate previously encountered items in the presence of PI was associated with the medial temporal lobe (MTL) and left ventrolateral prefrontal cortex (lVLPFC) activity (Oztekin et al., 2009). VLPFC is the critical region implicated in previous research that used neuroimaging techniques (Badre & Wagner, 2005, 2007), neuropsychological measurements (Jonides et al., 2002), and repetitive transcranial magnetic stimulation (Feredoes & Postle, 2010; Feredoes, Tononi, & Postle, 2006) investigating PI (see Jonides & Nee, 2006, for a review)(see Jonides, & Nee, 2006 for a review). This pattern was observed in studies examining PI with items differing in content (emotional, neutral, verbal, visual), suggesting a content-free role of lVLPFC in the control of PI (Badre & Wagner, 2005; Levens & Phelps, 2010; Mecklinger, Weber, Gunter, & Engle, 2003). Involvement of hippocampal and parahippocampal regions of MTL in episodic memory retrieval is well established as well (Diana, Yonelinas, & Ranganath, 2007; see Eichenbaum, Yonelinas, & Ranganath, 2007, for a review). One potential mechanism that aids the successful resolution of PI is strategic/controlled episodic retrieval (Badre & Wagner, 2005; Oztekin et al., 2009; Öztekin & McElree, 2007). This explanation has been supported by previous research that noted greater MTL activation during memory retrieval in the presence of PI compared to retrieval that was not subjected to PI. For instance, Öztekin, Curtis and McElree (2009) used a behavioral procedure that controlled the build up of PI via both semantic similarity of the items (i.e., the release from PI paradigm) and also by manipulating the recency of the probes as in the RP task. Their results indicated that resolving interference, regardless of semantic or episodic, required both MTL and lVLPFC regions.

Finally, Öztekin and Badre (2011) used the rfPI paradigm to test the effects of PI on the encoding and retrieval of semantically related words. They showed that during retrieval anterior part of left VLPFC (aVLPFC) varied as a function of PI, crucially, it was the mediator of the PI effects on behavioral measurements. Memory performance was less affected with higher aVLPFC activation. They also observed an increase in parahippocampal region activity in response to increasing levels of PI. Accordingly, in our task we also expect lVLPFC and MTL regions to respond to PI. We further hypothesize different neural activity patterns in these sites for emotion and neutral trials. Specifically, we do not foresee the involvement of hippocampal and parahippocampal regions of MTL in the processes that aid the recovery of PI for emotion trials. The rationale for this comes from the findings showing enhanced item-emotion bindings for negative and highly arousing memoranda compared to neutral memoranda (see Yonelinas & Ritchey, 2015, for a review). If item-specific details are enhanced for emotion items, retrieval of these specific details should not be as effortful for emotion trials compared to neutral trials. On the other hand, participants might need cognitive control to guide selective memory retrieval focusing on relevant and more effective cues while ignoring the shared cues between items for both types of stimuli. Optimal retrieval strategy in this task is to ignore the shared cue between the items and retrieve other features that might increase the probability of successful retrieval.

To summarize, the main goal of this study is to show how interference in working memory leads to forgetting of neutral and emotional information both behaviorally and neurally. There is mounting evidence suggesting that emotional memories are more resistant to forgetting compared to their non-emotional counterparts (reviewed in Yonelinas & Ritchey, 2015). We aimed to examine how PI evolves and how it is attenuated for emotional material. In our previous work (Mızrak & Öztekin, 2016), we modeled the full time course of retrieval and were able to show that PI for emotional items accumulated differently than neutral items. Here, we use the diffusion model to obtain estimates of cognitive processes that are involved in the recognition decision (Ratcliff & Starns, 2013; White & Poldrack, 2014, see Method section for detailed description of the model).We also examine neural activity during retrieval of emotion and neutral trials and how they change in response to increasing PI levels, which allows us to further relate the diffusion model parameters with neural responses from our target regions of interests.

3.3. Method

*Participants***.** Twenty-one healthy adults (9 female) participated in this study. Data from two participants (one male, one female) were excluded due to excessive motion in the scanner. Participants had normal or corrected to normal vision and were screened for medical conditions that could contradict with MRI protocols. Participants gave written consent and were compensated for their participation.

*Stimuli selection***.** The stimuli consisted of 2 neutral (Kitchen utensils and Furniture) and 2 emotional (Fear and Disgust) categories of images. Some of the emotional images were chosen from International Affective Picture System (IAPS) database (Lang et al., 2005). The rest was selected from Google images and rated by 18 different participants that did not participate in the actual experiment (See supplemental material 3 for the detailed description of stimuli selection).

Experimental design. The experiment consisted of two sessions of six runs that took place on two different days. Each run contained 36 experimental trials. In each trial, participants first studied five-items in the encoding phase, solved three math problems, and then responded to a recognition probe. For each probe, participants were asked to indicate whether it appeared in the encoding phase (See Fig. 1 for the procedure).

The release from PI paradigm was employed in which images from the same category were presented for three consecutive trials. The three consecutive trials from the same category formed a list and the first trial within this triplet will be referred to as List 1, the second as List 2, and the third as List 3. In total, there were 432 (216 x 2) trials (108 trials from each category), 144 lists (36 lists from each category). There were equal numbers of targets (i.e., probes that were shown in the encoding phase) and lures (i.e., probes that were not shown in the encoding phase) for each category. The probes of the three trials within one list were randomly chosen to be a target or a lure (e.g., it was possible that all three trials were lures or targets). After the presentation of one list (i.e., three trials) consisting of same category images, the next trial employed stimuli from another category.

Lures were drawn from members of the same category as the studied items. Both targets and lures were recycled throughout the whole experimental session with the restriction that neither had been presented within the last three runs before it was repeated. That is, all images used in the first three runs were novel, and in the following runs, images that were presented in the previous three runs were shown.

The order of lists from different categories was randomized in such a way that the number of switches between different categories was identical. This was done to make sure there was almost equal number of switches between different categories so that neutral to neutral or emotional to emotional switches happened equally often.

Procedure. Each trial proceeded with the presentation of a 5-image study list sequentially for 1200ms each. Following the fifth image, participants solved three math problems consisting of addition or subtraction of two randomly selected two-digit numbers. Participants indicated whether the solution presented next to the math problem was accurate by pressing either the middle or index finger on the button box. Participants had 4000ms to respond to each math problem. Following the third math problem, participants were presented with a test image and indicated whether the image was a member of the current study list. Participants had 2000 ms to respond to the test probe. The inter-trial interval consisted of presentation of a fixation cross on the center of the screen for fixed duration (12000ms).

fMRI protocol. Scanning was performed on a Siemens 3T Magnetom TRIO MRI system with a 32 channel head coil in the National Magnetic Resonance Imaging Center at Bilkent University. Functional images were acquired over six runs in each session using a gradient echo planar imaging (EPI) sequence (TR = 2000 ms, TE = 30 ms, flip angle = 90°, FOV = 192 mm, 34 interleaved axial slices, voxel size = 3 mm × 3 mm × 3 mm with 0.3 mm interslice gap). After each session, high-resolution T1-weighed (MP-RAGE) anatomical images were acquired. Participants' head motion was restricted by applying padding around the head. The experiment was presented on a screen via a projector which then was reflected to the mirror that was attached to the head coil. Participants gave their responses with their right hand through a standard four-button response pad.

Pre-processing. SPM8 (http://www.fil.ion.ucl.ac.uk/spm/software/spm8/) was used to process images which included slice timing correction, realignment, normalization, and smoothing. Functional images were first corrected for differences in slice acquisition timing by resampling all slices in time to match the first slice, then realigned for motion correction and resliced. Resliced images were then normalized to MNI stereotaxic space using a 12-parameter affine transformation along with a non-linear transformation using cosine basis functions. Functional images were then smoothed with an 8-mm FWHM isotropic Gaussian kernel. Image data quality was assured via visual inspection and runs in which there were excessive motion (> 3 mm) were excluded from the analysis.

fMRI data analysis. A general linear model was built in SPM8 by generating separate regressors for each condition [separate encoding for each of the three lists for emotion and neutral, distractor period (collapsed across lists both for emotion and neutral trials), recognition probes for each of the three lists for emotion and neutral] and were modeled using a canonical hemodynamic response function and its temporal derivative. Data across runs within each session was modeled as one session with mean signal and scanner drift entered as covariates.

Percent BOLD signal change was assessed by averaging the time-series signal across the region of interests (ROIs) by using the MarsBaR region of interest toolbox for SPM8. We calculated integrated percent signal change (iPSC) by averaging the peak time point and the two adjacent time points to the peak (±1 TR) for each ROI. Medial temporal lobe ROIs (including amygdala) were defined using anatomical masks from the Automated Anatomical Labeling (AAL) database (Tzourio-Mazoyer et al., 2002). We used left and right parahippocampal gryus (PHg) and hippocampus ROIs for MTL. For sub regions of left VLPFC ROIs we assessed pre-defined anatomical masks which were used in Öztekin and Badre (2011) study to test functionally dissociable roles of anterior and mid parts of leftVLPFC.

Modeling recognition decision. The diffusion model was used to model responses and corresponding response times (Ratcliff, 1978; Ratcliff & Rouder, 1998; Ratcliff, Smith, Brown, & McKoon, 2016). We employed the four parameters Wiener model (e.g., Wabersich & Vandekerckhove, 2014) augmented by a drift criterion parameter (Ratcliff & McKoon, 2008). In our task, the lower decision bound corresponded to “new” responses (i.e., when the participant decided that the probe is novel); the upper bound corresponded to “old” responses (i.e., when the participant decided that the probe belonged to the given study list). Across time evidence is accumulated in a noisy (diffusion) process. If the evidence hits one of the two response boundaries the corresponding response is given. The first parameter, the drift rate, v, with which the accumulation process approaches one of the boundaries given a specific stimulus class (i.e., old or new probe), is a measure of the strength of evidence resulting from the retrieval processes. In our model, the same drift rate is used for old and new probes with the only difference being their sign (positive for old probes and negative for new probes). The larger the absolute value of the drift rate the stronger the mnemonic evidence in that condition. The position of the response boundaries is captured by the boundary separation parameter *a*. The larger the *a,* the more evidence is necessary until a decision is made (i.e., the decision maker is more cautious).

Evidence starts to accumulate from a starting point which might be closer or at equal distance to one of the boundaries. This response bias parameter, *z*, measures the bias participants have towards one of the boundaries, with $z = 0.5$ indicating no bias. Participants might be biased towards an old response (if $z < 0.5$) in which case giving an old response would require less evidence. Note that response bias shifts do not result in a change in the drift rate.

A different possibility to account for response bias is via the drift criterion parameter, d_c. As mentioned above, we initially assume that the absolute value of the drift rate is identical for old and new probes. The drift criterion removes this assumption. More specifically, the value of the drift criterion is added to both drift rates, the positive drift for new items and the negative drift rate for old items and thereby represents a symmetric shift in the available evidence (while not affecting its absolute magnitude). Values above zero indicate that in a given condition the evidence is shifted towards old probes whereas values below one indicate a shift towards new probes (Ratcliff, & McKoon, 2008). Finally, our model also included a non-decision parameter, t_0, to capture all processes not related to the mnemonic evidence provided by the items such as encoding and motor processes.

Studies examining the impact of emotion on recognition memory suggest that emotion has an impact on the response criterion, traditionally measured within signal detection framework (Macmillan & Creelman, 2005). By separating response bias and drift criterion, we can a) evaluate the retrieval strategies participants adapt to changing levels of PI and b) investigate whether emotion has an impact on the response bias or drift criterion.

We modeled the data in a hierarchical Bayesian fashion on the responses made for each trial and corresponding response times. For each of the five parameters (i.e., v, a, z, d_c, t_0) we estimated six hyperparameters to describe the full design (i.e., emotion conditions [emotion & neutral] and list [List 1, List 2, & List 3]). More specifically, we estimated the model using treatment contrasts separately for emotional and neutral stimuli such that List 1 was the intercept plus two parameters quantifying the difference of List 2 and List 3 from the intercept (i.e., we had three parameters, one intercept plus two increments, for both emotional and neutral items). In total, the model had 30 hyperparameters.

To model the hierarchical structure we employed the latent-trait approach of Klauer (2010). To account for the individual variability we estimated normally distributed zero-centered deviations for each participant and each of the 30 parameters that were added to the corresponding hyperparameter. In addition, we estimated the full variance-covariance matrix among the individual deviations, allowing them to be correlated across parameters. This model is thereby analogous to the maximal random effects structure of regression models (Barr, Levy, Scheepers, & Tily, 2013).

Following Gelman et al. (2014) we used weakly informative priors for all parameters to facilitate convergence of the model. For the hyperparameters of the drift rate, v, we used Cauchy (1, 2) [location, scale] priors for the intercepts and Cauchy (0, 2) priors for the increments. For the hyperparameters of the boundary separation, a, we used Cauchy (1, 2) priors for the intercepts and Cauchy (0, 1) priors for the increments. For the hyperparameters of the drift criterion, d_c, we used Cauchy (0, 1) priors for both intercepts and increments. For the hyperparameters of the starting point, *b*, we used Normal (0.5, 0.5) [mean, standard deviation] priors for the intercepts and Normal (0, 0.5) priors for the increments. For the hyperparameters of the non-decision time, t_0, we used Normal (0.2, 0.2) priors for the intercept and Normal (0, 0.2) priors for the increments. The prior for individual deviations was a Multivariate Normal distribution with mean 0 and standard deviation of 1. The variance-covariance matrix was split into priors for the variances with Cauchy (0, 4) and correlation matrix with LKJ (1) using the method of Lewandowski, Kurowicka, and Joe (2009).

The model was estimated using a version of Hamiltonian-Monte-Carlo using Stan (Carpenter et al., in press). We estimated a total of 1000 samples in 6 separate chains. The first 250 samples were discarded as warmup samples and we retained every second sample from the remaining samples. Model convergence was assessed based on $\hat{R}$ (all $\hat{R} < 1.01$) and visual assessments of chain convergence. Figures S3.1 and S3.2 show posterior predictive checks which compare data simulated from the model to the empirical data and demonstrate that the RT distributions and response proportions for each conditions and participant are overall well recovered (see Supplemental Material 3).

3.4. Results

Behavioral Data

Memory performance worsened as a function of PI. Participants were either less accurate and/or slower at recognizing items in high PI trials (List 3) compared to low PI trials (List 1) as shown in Figure 3.1.

Accuracy. We assessed impact of PI on memory performance by looking at correctly recognized old items (hits) and falsely recognized new items (false alarms). To do so, we conducted a 2 (stimulus type [Emotion vs. Neutral]) x 3 (PI [List 1 = No PI, List 2 = medium PI, List 3 = high PI]) analysis of variance (ANOVA) on both hits and false alarm rates. For false alarms, we found a main effect of PI: *F*(1.50, 27.07) = 6.16, *p* = .01, η_g^2 = .04 such that PI increased the false alarm rate. We also observed an interaction of type of stimuli and PI, *F*(1.42, 25.55)=4.48, *p* = .03, η_g^2 = .03, showing that this effect occured for neutral stimuli only. For neutral trials, retrieval success decreased from List 1 to List 3, whereas emotion trials were affected by PI increase in lower amounts. We also observed a main effect of stimulus type (*F*(1,18) = 5.41, *p* = .03, η_g^2 = .06) showing that participants made more false alarms to neutral trials than to emotion trials.

For the hit rate, we found no main effect of PI nor an interaction (main effect of PI: *F(*1.91,34) = 1.44, *p* = .25, η_g^2 = .009, interaction between PI and stimulus type: *F(*1.98,35.68) = 0.27, *p* = .77, η_g^2 = .001), change in PI levels did not decrease the hit rates. Emotion improved hit rates at all levels of PI (main effect of stimulus type: *F(*1,18), p < .001, η_g^2 = .20).

Response times. We assessed response times (shown in Figure S4) to correct rejections and hits with a 2 (stimulus type [Emotion vs. Neutral]) x 3 (PI [List 1 = No PI, List 2 = medium PI, List 3 = high PI]) ANOVA. For hits, we did not observe a main effect of PI [*F*(1.48, 26.70) = 2.66, *p* = .10, η_g^2 = .04]. However, there was a significant interaction between list and stimulus type showing that responses slowed as a function of PI to recover from the interference from irrelevant material only for emotion trials [*F*(1.76,31.62) = 45.7, *p* = .05, η_g^2 = .04]. This effect exhibited itself as a linear decrease in the RTs to hits from List 1 to List 3 (estimated slope = 0.03, t(70.06) = 3.02, *p* < .01) whereas we did not see such an increase for neutral trials (*p* = .85). Note that, participants exhibited faster responses to hits for emotion trials (main effect of stimulus type: $F(1,18) = 45.7$, $p < .001$, $\eta_g^2 = .05$).

Next, we examined response times to correct rejections and found that participants got slower with increasing levels of PI (main effect of list: $F(1.95, 35.14) = 4.78$, $p = .02$, $\eta_g^2 = .005$). Participants were also slower for neutral compared to emotional stimuli (main effect of stimulus type: $F(1, 18) = 29.07$, $p < .001$, $\eta_g^2 = .03$). We did not observe an interaction between list and stimulus type ($p = .52$), however post hoc comparisons showed that there was a linear increase in response times for neutral trials (estimated slope = 0.032, $t(71.6) = 1.33$, $p < 0.2$) but not for emotion trials ($p = .18$).

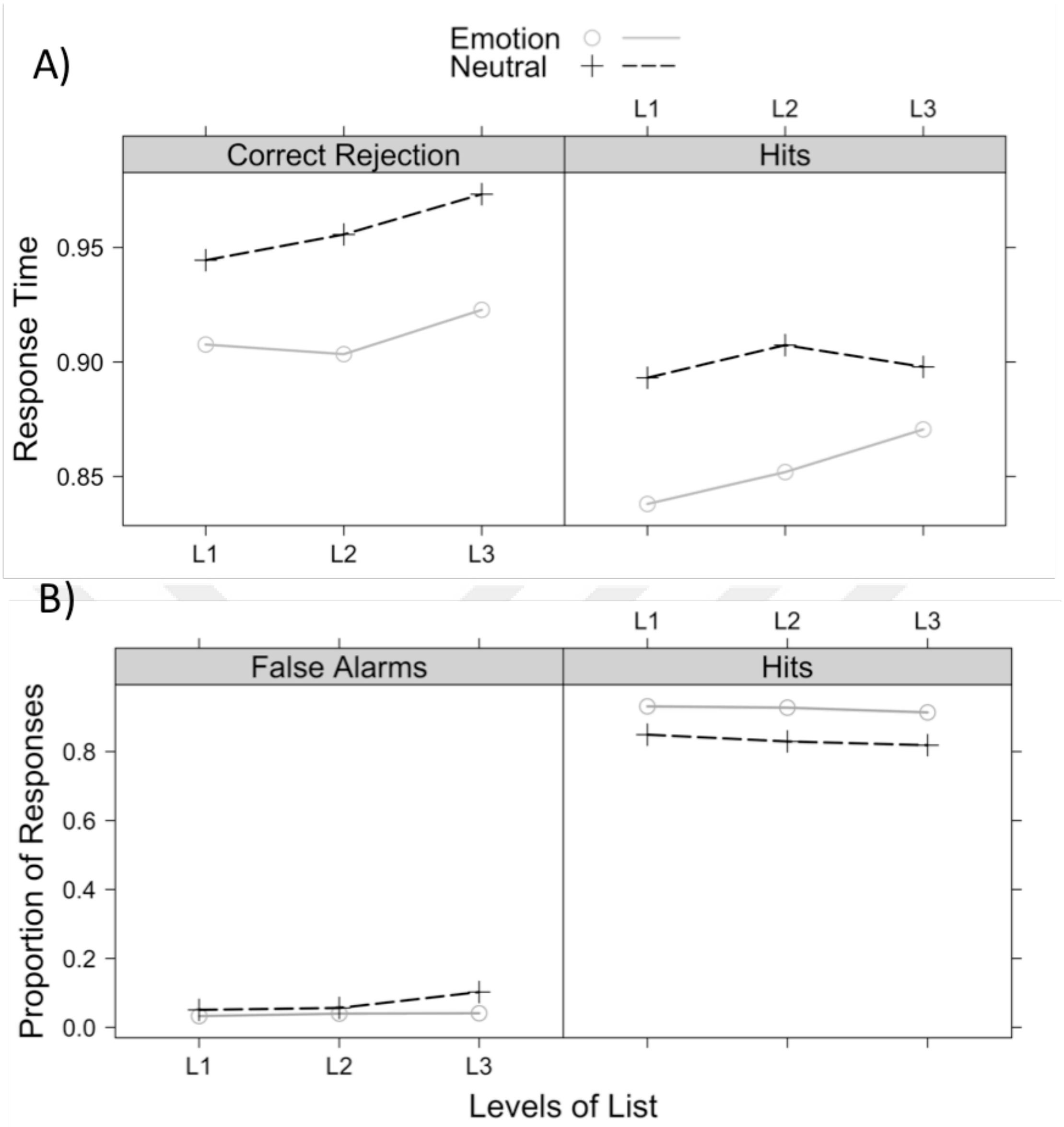

Figure 3.1. A) Mean response times to correct rejections and hits for emotion and neutral trials. B) Mean proportion of responses as false alarms and hits for emotion and neutral trials. L1 = List 1, L2 = List 2, L3 = List 3.

Hierarchical Bayesian diffusion model analysis

Diffusion models have often been used to model recognition memory decisions (e.g., Ratcliff & Starns, 2013; Starns, Ratcliff, & McKoon, 2012; White & Poldrack, 2014). Consistent with prior work showing that PI slowed down the retrieval speed (Öztekin & McElree, 2007), we hypothesized that PI would reduce the efficiency of evidence accumulation (i.e., decrease the drift rate) due to reduced memory signal quality with increasing similarity between memoranda.

Results confirmed this prediction. As can be seen in Figure 3.2 (panel "Drift Rate") the drift rate decreased linearly for neutral trials with increasing PI levels (List 1 vs. List 2: p_B = .01; List 2 vs. List 3: p_B = .01)[5]. However, while descriptively such a pattern was also obtained for emotion trials, the evidence for it was comparatively weak. There was no significant decrease from List 1 to List 2 (p_B = .27) and neither from List 2 to List 3 (p_B = .11). Only when comparing List 1 and List 3 we found a significant decrease in drift rate (p_B = .027). When comparing the decreases in drift rate due to PI across emotion and neutral items there was little evidence for a differential pattern (smallest p_B = .062). This suggests that PI affected both stimulus types, although there was a tendency that this decrease was less pronounced for emotion trials.

In addition to the effect on drift rate, PI also affected the drift criterion. Figure 2 shows that, overall, there is a tendency for a conservative bias (i.e., $d_c < 0$). We can furthermore see that neutral and emotional items seem to exhibit an almost opposite pattern. For neutral items it appears that the drift criterion becomes more extreme (i.e., smaller) from List 1 to List 2, but then decreases again for List 3, whereas for emotional items the drift criterion appears to increase from List 1 to List 2 (however, none of these comparisons reaches significance, smallest p_B = .13) and becomes evidently more extreme for List 3 (p_B = .041). In support of this differential pattern, the increase for neutral items from List 2 to List 3 tends to be larger than the decrease for emotional items from List 2 to List 3 (p_B = .022).

Figure 3.2 further suggests a similar pattern for the response bias. For emotional items, we again see a (non-significant) decrease in bias from List 1 to List

---

[5] p_B-values reported in this manuscript are calculated from the posterior distributions of the hyperparameters and based on the probability that specific parameter or difference of parameters is smaller or larger than 0 (i.e., values > .5 are subtracted from 1).

2 (p_B = .17) which increases again towards List 3 (p_B = .021). For neutral items, however, there is a (non-significant) increase from List 1 to List 2 and further to List 3 (smallest p_B = .10). Furthermore, we tested whether the difference due to PI in response bias were different for emotion and neutral trials, however, they were not (smallest p_B = .06). Overall, participants showed a response bias towards old responses.

Non-decision time and boundary separation estimates were comparable between emotion and neutral trials. We found an effect of PI on non-decision time for neutral stimuli, which decreased from List 1 to List 2 (p_B = .016) with List 3 in between the two. We did not observe such a strong effect for emotion trials (smallest p_B = .052). Boundary separation for neutral stimuli increased from List 1 to List 2 (p_B = .042) while it stayed at similar levels from List 2 to List 3.

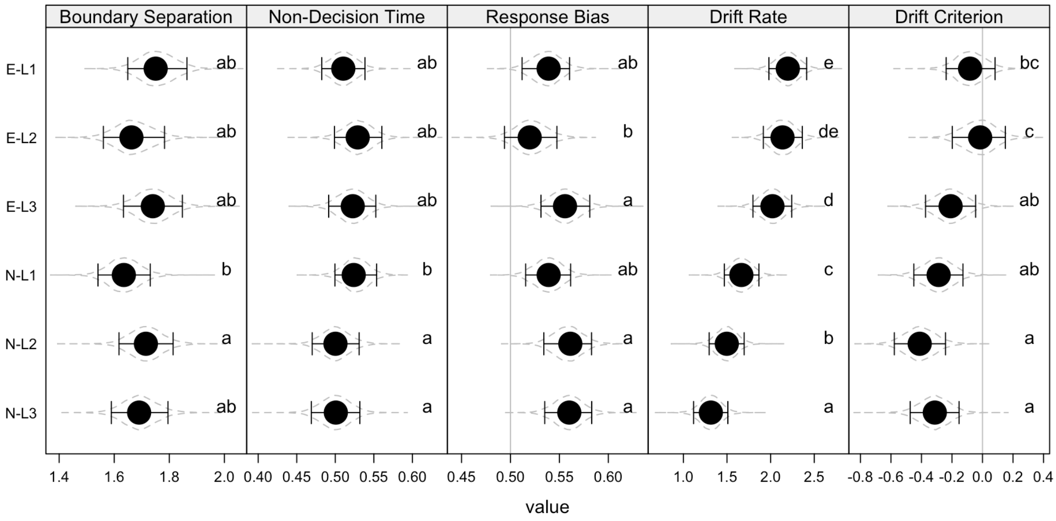

Figure 3.2. Parameter estimates from the diffusion model. E = Emotion, N = Neutral, L1 = List 1, L2 = List 2, L3 = List 3. The points show the posterior modes, the error bars the 90% highest-posterior density regions, and the gray lines the (mirrored) density estimates of the full posterior. The letters in each plot represent a compact letter display (CLD; Piepho, 2004) presentation of the difference between conditions. Conditions which do not share a letter within one plot differ significantly from each other with p_B < .05.

Neuroimaging Data

Effects of PI on the recognition of emotion and neutral stimuli.

Prior work on cognitive control of PI in working memory showed that left ventrolateral prefrontal cortex (lVLPFC) (reviewed in Jonides & Nee, 2006) and specific regions within MTL are critical for resolving PI (Öztekin & Badre, 2011; Oztekin et al., 2009). Additionally, research suggested that sub-regions of lVLPFC, namely mid-VLPFC and anterior-VLPFC (aVLPFC), might play different roles in this process. Öztekin and Badre (2011) showed that both lVLPFC and the anterior part as well as parahippocampal regions of MTL aided the control of PI in the release from PI paradigm. aVLPFC activation correlated with the response time increase from low PI (List 1) to high PI (List 2 & 3) conditions, mediating the impact of interference on behavioral measures. They also showed the parahippocampal cortex activation increased with the build up of PI for correctly recognized target (i.e., old) items. Following this, we first assessed the integrated percent signal change (iPSC) in the lVLPFC, testing both aVLPFC and mid-VLPFC, and hippocampal and parahippocampal regions (See method for the definition of regions). We conducted a 2 (stimulus type [Emotion vs. Neutral]) x 3 (Lists) Analysis of Variance (ANOVA) on iPSC for correctly recognized target items of left and right PHg revealing a main effect of condition for both regions: activation was greater for emotion trials in PHg [left; F (1,18) = 5.62, p < .03, η_g^2= .05, right: F(1, 18) = 4.0, p = .06, η_g^2 = .04] while there was no reliable impact of PI [left; F (1.69, 30.45) = 1.63, p = .22, η_g^2 = .1, right: F(1.85, 33.23) = 2.36, p = .11, η_g^2 = .02] and no interaction between PI and stimulus type [left; F (1.88,33.75) = 1.99, p = .15, η_g^2 = 0.01, right: F(1.74, 31.36) = 1.96, p = .16, η_g^2 = .01]. We nevertheless inspected the pairwise comparisons of emotion and neutral trials for each list and observed that emotion only had greater activation in both left and right PHG for List 1 (i.e., in the absence of PI) compared to neutral trials (Left:L1: $p < .01$, Right:L1: $p < .01$) and this difference dissipated with increasing levels of PI (smallest $p = .14$). Post hoc comparisons also indicated that left and right PHG activation linearly increased as a function of lists ([Left : estimated slope = .08, $t(69.21) = 2.258$, $p < .03$, Right: estimated slope = .09, $t(71.83) = 2.61$, $p < .02$]) for neutral trials. We did not observe such a pattern for emotion items in parahippocampal regions, activation levels did not change as a function of PI levels (smallest $p = .71$).

Whereas we replicate the linear increase in PHg (both left and right), activation in VLPFC regions was not significantly affected by the build up of PI. For aVLPFC we estimated the same ANOVA and again did not see a main effect of PI [F (1.68, 30.15) = 0.55, $p = .55$, $\eta_g^2 = .005$]. We also did not observe an interaction between PI and stimulus type [F (1.76, 31.76) = 0.65, $p = .51$, $\eta_g^2 = .005$]. However, aVLPFC activation was reliably greater for emotion compared to neutral [F (1, 18) = 4.03, $p = .06$, $n_g{}^2 = .04$]. Inspection of all pairwise comparisons revealed that there was no effect of stimulus type for List 1 and List 2 (L1: $t(43.69) = 1.58$, $p = .12$, L2: $t(43.69) = 0.81$, $p = .41$), but it was observed in the high PI condition (L3: $t(43.69) = 2.113$, $p = .04$]). As shown in Figure 3.3 we also observed an increase from List 2 to List 3 for emotion trials but this difference was not statistically reliable (L1: $M = 0.13$, L3: $M = 0.18$).

Öztekin and Badre (2011) showed a negative correlation between aVLPFC activation and the corresponding decline in memory performance measured by RT increase as a result of PI. Following their logic and the fact that we found a small (albeit not significant) increase for emotion trials, we examined the relationship between the difference in aVLPFC activation between lists and RT increase for emotion trials. We found a significant correlation between the two measures showing that as the aVLPFC activation difference between lists increased, RT difference decreased [L1-L3 difference: *r*(17) = -.425, *p* = .07, L2-L3 difference: *r*(17) = -.528, *p* = .02]. The more aVLPFC was engaged, the less impact PI had on memory performance (see Figure 3.4).

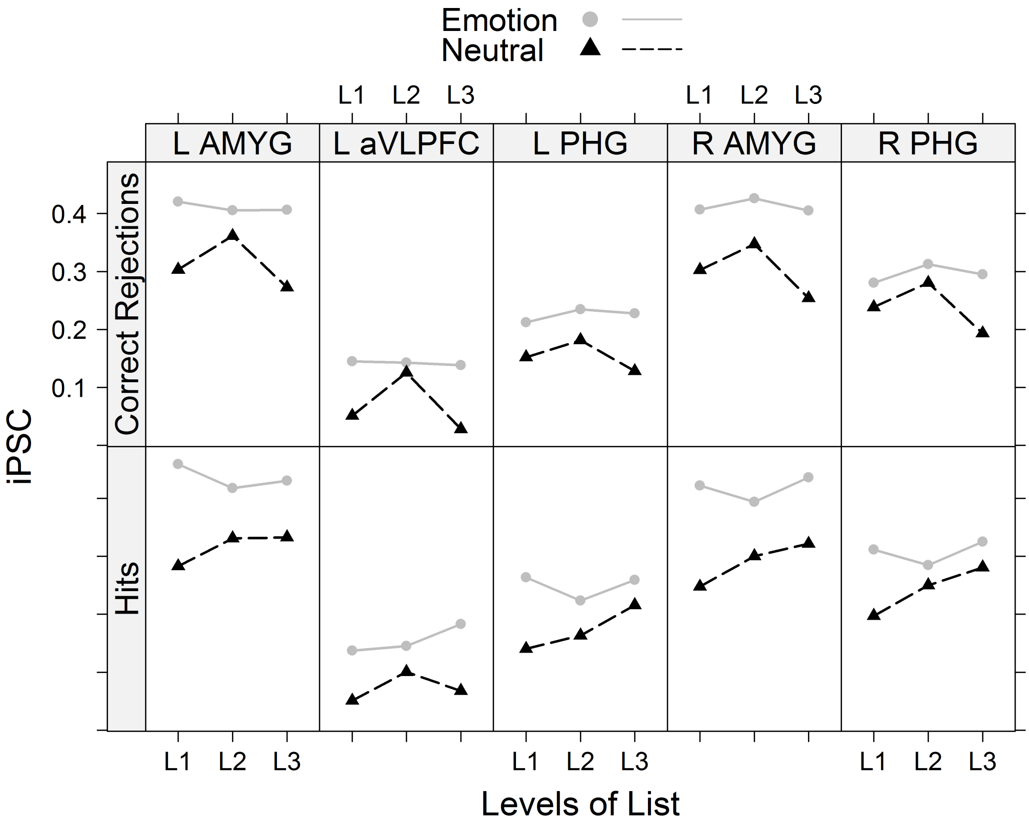

Figure 3.3. Integrated percent signal change (iPSC) values for List [List 1, List 2, List 3] * Stimulus type [Emotion, Neutral] * Recognition response [Hits, Correct Rejections]. Hits correspond to correctly recognized target probes, correct rejections are correct responses to lures. AMYG = amygdala, aVLPFC = anterior ventrolateral prefrontal cortex, PHg = parahippocampal regions, L = left, R = right.

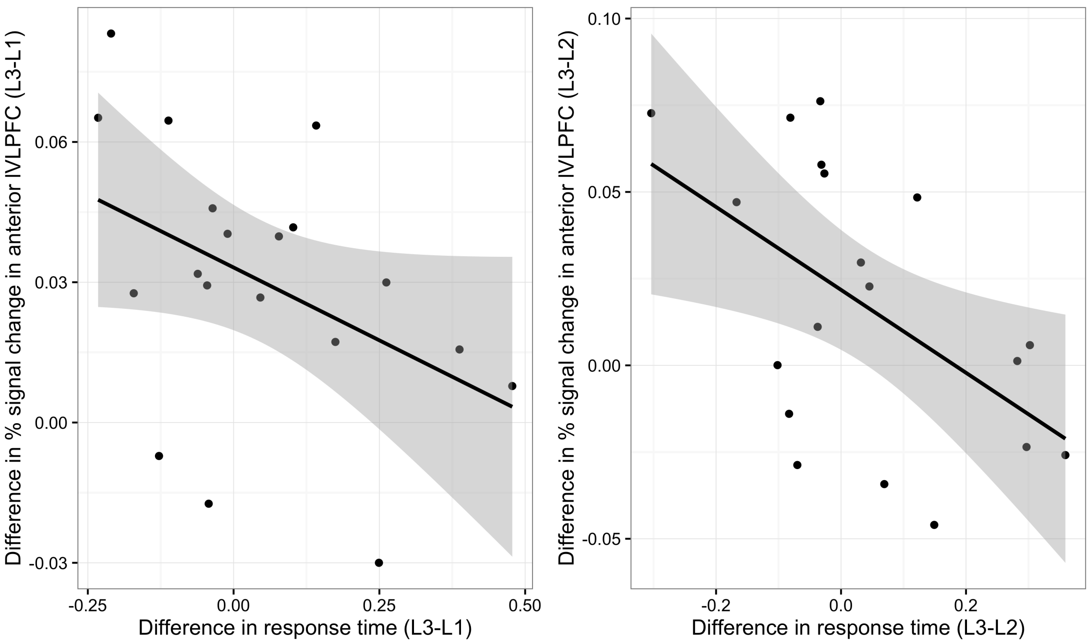

Figure 3.4. Left panel shows the correlation between the corresponding increase in response times and the integrated % signal change in aVLPFC region from List 1 to List 3. Right panel exhibits the relationship between aVLPFC and response time increase from List 2 to List 3.

Differences between emotion and neutral stimuli at retrieval in amygdala

Studies which investigated emotion-memory interactions consistently demonstrated higher amygdala activation for emotional memoranda than neutral memoranda (Florin Dolcos, LaBar, & Cabeza, 2004). We investigated possible differences in left and right amygdala. We examined the amygdala response to hits and correct rejections. For both types of responses, amygdala (left and right) activation was stronger to emotional memoranda [Hits:Left: F (1, 18) = 10.38, p = .005, $n_g{}^2$ = .07; Hits:Right: F(1, 18) = 10.21, $p =$ .005, η_g^2 = .09; CRs:Left: F(1, 18) = 3.67, p = .07, η_g^2 = .04; CRs:Right: F(1, 18) = 7.95, p = .01, η_g^2 = .07].

Linking latent processes with neural measures

Our ROI analysis replicated Öztekin & Badre's (2011) results showing a linear increase in PHg activation for hits to neutral items as a function of PI. As we also observe a linear decrease in drift rate from List 1 to List 3, we investigated the relationship between PHg and drift rate changes for hit responses. To adequately account for the repeated-measures nature of the data (each participant provided three drift-rate estimates and three corresponding activations per stimulus type for hits), we conducted a linear mixed model (LMM; Baayen, Davidson, & Bates, 2008) analysis. More specifically, we estimated a LMM on the individual posterior mean drift rates with fixed effects for PHg (mean across PHg-L and PHg-R), stimulus type, and their interaction, as well as a by-participant random intercept and by-participant random slopes for all fixed effects (this model thereby implements the maximal random effects structure; Barr et al., 2013)[6]. To evaluate the fixed effects we employed the Kenward-Roger approximation via the methods implemented in afex (Singmann, Bolker, & Westfall, 2015).

The model showed a significant effect of PHg [$F(1, 53.43) = 7.41$, $p = .01$], a significant effect of stimulus type, [$F(1, 37.89) = 30.75$, $p < .001$], as well as a marginally significant interaction between PHg and stimulus type [$F(1, 33.45) = 3.58$, $p = .07$]. Figure 3.5 displays the interaction and shows that PHg activation is a significant predictor for the drift rate for neutral items ($\beta = -0.70$, 95%-CI [-1.13, -0.26]), but not for emotional items ($\beta = -0.15$, 95%-CI [-0.56, 0.26]). This suggests that the decrease in drift rate due to PI for neutral items was accompanied by corresponding increase in PHg activation. Increased PHG activity presumably indicates that participants tried to recover from PI by accessing detailed item information.[7]

---

[6] To overcome convergence problems, we removed the correlation between random slopes.

[7] Notwithstanding that, we did not observe such a relationship for sub regions of VLPFC and amygdala (largest $F = 1.35$).

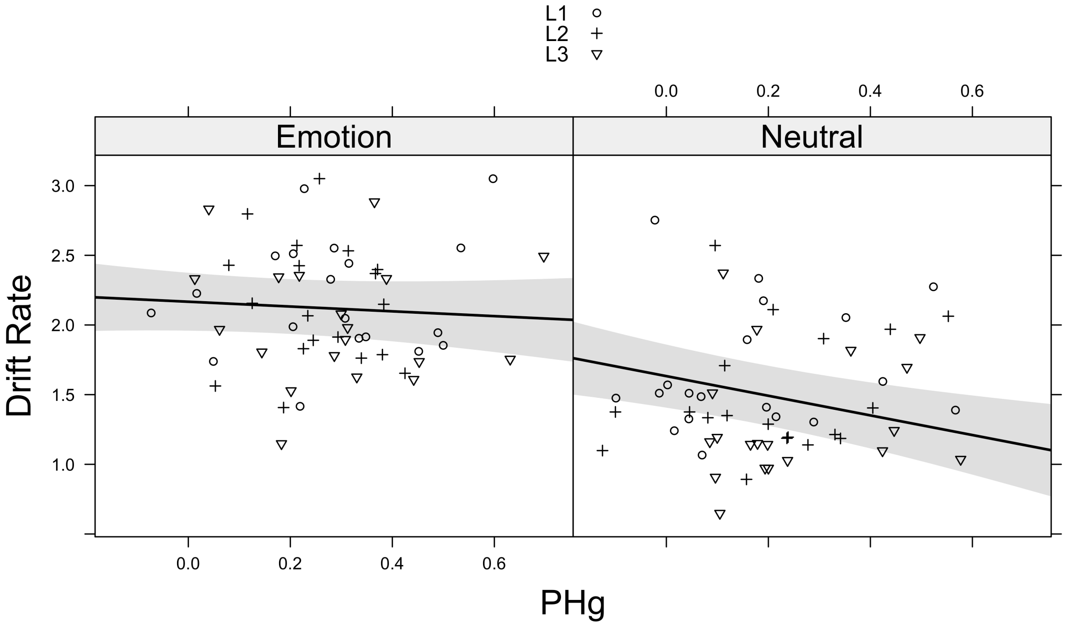

Figure 3.5. The interaction between parahippocampal gyrus activation (PHg) and drift rate for neutral (β = -0.70, 95%-CI [-1.13, -0.26]) and emotional items (β = -0.15, 95%-CI [-0.56, 0.26]). Shaded area shows 95%-confidence bands from the LMM.

3.5. Discussion

The present study aimed to provide new insights into the relationship between emotion and forgetting due to PI from working memory. We were specifically interested in differences in a) the build up of PI and b) recovery from PI between emotional and neutral stimuli. To this end, we used a task that progressively induced PI by presenting categorically similar items over three trials. After three trials, the category was changed which led to a release from PI. Firstly, we employed a formal measurement model, the diffusion model, that allowed us to estimate and compare how the built up of PI affected latent cognitive processes of recognition memory decisions of emotional and neutral material. Second, we tracked the neural responses to PI during retrieval of emotional and neutral memoranda in regions that have been previously related to overcoming PI.

Effects of PI on latent processes

We showed that the rate of evidence accumulation (drift rate) decreased as a function of PI for both neutral and emotional stimuli. Drift rate here represents the quality of the match between probe and memoranda (Ratcliff & McKoon, 2008), it is a measure of the strength of evidence. Items with higher memory strength would have a higher drift rate. Our findings showed that PI decreased the drift rate (i.e., lowered the mnemonic quality/degree of match). It has been suggested that one of the factors that affects the strength of evidence in memory retrieval might be the degree of difficulty of recovering the evidence (Badre, Lebrecht, Pagliaccio, Long, & Scimeca, 2014). In our task, the shared feature (category) elicits a level of match between the probe and memoranda, contaminating the quality of the match and adding noise to the evidence making the recovery of evidence harder.

Our results indicate that the PI-related decrease in the quality of the probe-memoranda match is similar for emotion and neutral stimuli. This is an important finding showing that PI in WM affects memory performance for emotion and neutral stimuli in a similar fashion. It is noteworthy to state that the decline in drift rate tended to be less pronounced for emotion stimuli such that medium levels of PI did not lead to an effective decline in drift rate. This is consistent with our previous findings in Mizrak and Öztekin (2016) study showing that PI build up was slower and less for emotion trials compared to neutral trials. We discuss potential explanations for this finding below.

Another notable finding from our study was the shifts in response level bias and memory level bias (drift criterion). For neutral stimuli, participants became more biased from List 1 to List 2 (towards responding “old” and a complementary shift for the drift criterion towards evidence for new responses) which tended to decrease again from List 2 to List 3. For the emotional stimuli a similar pattern, but in the opposite direction (i.e., less bias from List 1 to List 2 and then more bias towards List 3), was observed. It has been suggested that an individual can adjust her/his drift criterion depending on the context or the relevant task goal or even the expected value of the outcome (Scimeca, Katzman, & Badre, 2016). Our findings indicate that participants tried to react to the increasing PI by adapting their response and memory strategies. As this did not help in solving the task they reverted the initial adaptations. Note that the possible alternative explanation of parameter trade-offs is not supported by the data as we did not find corresponding correlations among the individual level parameter estimates.

Neural mechanisms that overcome PI for emotion and neutral stimuli

Neutral stimuli. For neutral stimuli, increasing levels of PI led to a gradual decline in the rate of evidence accumulation. This decrease was however accompanied by an increase in activation in parahippocampal regions (PHg) for hits (i.e., “old” responses to targets). This suggests that participants tried to counteract the effect of PI by increasing memory retrieval. This is in line with previous work showing that MTL regions are involved in successful episodic retrieval (e.g., Dobbins, Rice, Wagner, & Schacter, 2003). Additionally, it has been shown that MTL activation during encoding predicted memory success (e.g., Davachi, Mitchell, & Wagner, 2003; Paller & Wagner, 2002; Staresina & Davachi, 2006). More recently, MTL activation has been observed during retrieval from WM (Oztekin, Davachi, & McElree, 2010; Öztekin et al., 2008), also suggesting a relationship between memory accuracy and MTL activation. Notwithstanding, we did not detect changes in aVLPFC activation in response to PI. This finding is surprising as this region had often been suggested to mediate the detrimental effects of PI on memory performance

(Badre & Wagner, 2005; Jonides & Nee, 2006; Oztekin et al., 2009).

Emotional stimuli. For emotional stimuli we only observed a comparatively small decline in drift rate as a function of PI. In addition, this decrease was not accompanied by a corresponding PHg activation increase to hits. This finding, as explained in the introduction, was expected due to enhanced item-emotion bindings for emotional material eliminating the need to retrieve the weaker item features to recover from PI.

We expected to see aVLPFC activation related to PI effects. However, we could not link the PI related changes in latent processes for emotional material with aVLPFC response to PI. Nevertheless, we were able to replicate the findings from Öztekin and Badre (2011) showing that the PI related response time increase was lessened to the extent that the aVLPFC response increased. This suggests that modulated aVLPFC activation might have helped participants to recover from detrimental effects of PI.

Effect of Emotion on PI

It has been suggested that the release from PI task specifically necessitates controlled retrieval at the item level (Oztekin et al., 2009). Participants need to access more diagnostic information about the item on top of the category membership. These details might be less associated with the item than the category membership of the item which might necessitate controlled retrieval. One common finding from the emotion-recognition memory literature is that emotion impacts retrieval of qualitative information about the event from memory (often called *recollection*) rather than assessments of the overall memory strength of the items in memory (often called familiarity; Kensinger & Corkin, 2003; Ochsner, 2000; Pierce & Kensinger, 2011; Ritchey et al., 2008; Sharot & Yonelinas, 2008). Specifically, emotion enhances recollection of within-item features such as identifying details about the item rather than context (Chiu et al., 2013; Yonelinas & Ritchey, 2015). In our task, participants might have benefited from a focus on the central details (item features) in exchange for remembering peripheral details (context) for emotional stimuli. If emotional stimuli are activated in WM with other members of the same category, these more diagnostic features can be used as valid cues while making the recognition judgment and the invalid cue category membership will have less weight on the recognition judgment. This can explain the less pronounced PI-related decline in drift rate for emotion stimuli compared to neutral stimuli. Due to consecutive presentation, category becomes a strong retrieval cue for both emotion and neutral stimuli. In this case, the category cue will contaminate the quality of the match between probe and memoranda, but for emotional stimuli not as much as it does for neutral stimuli.

This explanation can also account for the different pattern found by Mizrak and Öztekin (2016) in the recent probe task. At the end of the retrieval time course, participants in their study showed worse performance for emotional compared to neutral stimuli. In the recent-probes task participants specifically need to retrieve contextual details to overcome PI whereas in the current rfPI task it is necessary to access diagnostic item details to recover from PI. As emotional stimuli enhance memory for the latter, but not for the former, worse performance due to PI in the recent probe task can be expected.

3.6. Conclusion

We provided an in-depth investigation of the relationship between emotion and forgetting in WM by modeling the latent processes that are involved in the recognition memory performance of emotion and neutral items in the presence of PI. We complemented our investigation by analyzing neural responses to changing levels of PI and related our findings from brain and behavior measurements. We show that PI-related drift rate decline manifests itself in a similar fashion for both stimulus types. However, the degree of decline was slighlty less for emotional stimuli. Ventrolateral prefrontal cortex but not parahippocampal regions of medial temporal lobe counteracted the effect of PI for emotional stimuli and the opposite neural pattern was observed for neutral stimuli. For both stimulus types participants adjusted their response and decision biases in response to PI. We suggest that while cognitive processes affected by PI are similar for both types of stimuli, due to differences in the memory representations, the quantitative impact of PI reflected on recognition performance is different for emotional and neutral stimuli. Here, we investigated PI from highly arousing and unpleasant memoranda. Future studies could further evaluate the impact of arousal by taking the arousal component off the stimuli and inducing arousal during retrieval in different ways.

THESIS DISCUSSION

This dissertation consists of three pieces of work focusing on the controlled retrieval of relevant information from working memory. Our main research question was; what affects the efficiency of controlled retrieval process? The first chapter examined the impact of working memory capacity on controlled serial search that is required to retrieve temporal order memory. The second and third chapter focused on how retrieval dynamics in the presence of proactive interference were altered by the type of stimuli (emotion vs neutral).

The goal of the first chapter was to understand the role of working memory capacity on the dynamics of temporal order memory retrieval. Our aim for this chapter was two-fold; a) to reveal whether individuals with high working memory capacity perform superior when the task demanded controlled retrieval from working memory is required b) whether group differences would be reflected in the retrieval speed, accuracy or both. Specifically, we tested high span and low span individuals with a relative judgment of recency paradigm, in which they were asked to retrieve the temporal information of the items presented and judge the recency of the test probes. We assessed high and low WMC individuals' retrieval accuracy and retrieval speed parameters which were estimated by fitting the accuracy and response latency data with the full time-course retrieval functions. The overall accuracy of temporal order information retrieval, measured by asymptotic accuracy and asymptote estimated derived from quantitative modeling of SAT functions, was lower for low span individuals than high span individuals except for the test probes containing the most recently studied item. This finding was consistent with many studies in the literature showing WMC measures correlating with the probability of retrieval and high span individuals having better accuracy measures in memory tasks compared to low span individuals. More important and novel finding of this chapter was the retrieval speed differences across the two groups. Specifically, low span individuals were considerably slower in engaging in the controlled serial memory search operations that access temporal order information from working memory. Findings from the first chapter point to the contention that WMC predicts the ability to search through items in memory which requires controlled retrieval. We concluded that efficiency of controlled retrieval, measured by the retrieval speed and retrieval accuracy, depends on one's ability to maintain representations in an accessible state; working memory capacity. Our findings contribute to a growing literature showing that WMC differences are also reflected in controlled retrieval processes in addition to allocation of controlled attention. These findings should be taken into account when designing training programs that target WMC improvements.

In the second chapter, we were interested in a different factor that might impact controlled retrieval from WM; affective content of the stimuli. There is a growing body of evidence showing that emotional memoranda are resistant to forgetting from long term memory compared to neutral memoranda (reviewed in Yonelinas & Ritchey, 2015). One way forgetting happens in working memory is when previously learned yet irrelevant memoranda interferes with remembering of relevant memoranda. Previous research showed that although working memory is constantly subject to proactive interference, we overcome forgetting via engaging with controlled retrieval operations (Öztekin, & McElree, 2007). Here, we manipulated the need for controlled retrieval by inducing proactive interference at specific trials. Critically, we manipulated the content of stimuli such that half of the trials consisted of neutral stimuli and the other half were of highly arousing stimuli presented in a modified version of probe recognition task. Our goal was to reveal how emotional and neutral PI build up early in retrieval and how they are resolved via episodic/controlled retrieval later in retrieval. To do so, we examined the full-time course of false alarm differences between the recent lures that elicited this episodic familiarity and non-recent lures which did not. The false alarm difference here is an indicator of the impact of PI such that higher false alarm difference means PI had a higher detrimental effect on memory performance. The rationale behind examining the time-course of false alarms comes from the finding that familiarity as a source of information becomes available earlier than detailed source information about the item. Such a pattern would exhibit itself as higher false alarm differences early in retrieval when participants mostly rely on familiarity of the item and a decline in false alarms later in retrieval. We found that PI build up was slower and less for emotional stimuli compared to neutral stimuli. However, even though PI was less pronounced for emotional information recovering from PI via controlled retrieval was less efficient. This chapter raised an important question: is there a fundamental difference between emotion and neutral memoranda in the way proactive interference leads to forgetting and how it is overcome via controlled retrieval operations?

The third chapter followed up on this question we were interested in the second chapter: how does emotion affect proactive interference build up and its resolution through controlled retrieval of relevant memoranda? We had two goals in this chapter. First, we wanted to show how PI affects latent processes contributing to recognition memory decision for emotional and neutral memoranda. Second, we were interested to reveal whether the latent processes affected by PI and brain regions involved in the recovery of PI were similar for different types of stimuli. We used the diffusion model to obtain estimates of PI effects on cognitive processes that are involved in the recognition decision (Ratcliff & Starns, 2013; White & Poldrack, 2014). Additionally, we assessed neural activity during retrieval of emotion and neutral trials and how they change in response to increasing PI levels, which allowed us to further relate the diffusion model parameters with neural responses from our target regions of interests. We show that PI-related decline was reflected in the drift rate in a similar fashion for both types of stimuli. However, the amount of the decline was relatively less for emotion stimuli. Recovering from emotional PI did not depend on medial temporal regions but required ventrolateral prefrontal cortex and this pattern was in the opposite direction for neutral PI. We concluded that due to the nature of differences in emotional and neutral stimuli, detrimental effects of PI reflected on recognition memory performance are to a different extent across types of stimuli. However, the PI related decline in evidence accumulation efficiency for both types of stimuli show that PI build up is reflected in the same latent processes. Additionally, anterior VLPFC responding to PI for emotional stimuli suggest that recovery from PI on this task depends on similar brain regions that had been shown previously with this task that is thought to be involved in controlled retrieval processes (Öztekin & Badre, 2011). Overall, we conclude that there are no fundamental differences between emotional and neutral information in the way PI leads to forgetting and how it is overcome. Yet, both chapters provided evidence towards a less pronounced PI build up for emotional material. Importantly, in this chapter we show that affective content of stimuli does not lead to less efficient controlled retrieval when to-be-retrieved information is item specific details and not contextual information.

Collectively, findings from the three chapters indicate that efficiency of controlled retrieval from working memory is affected by working memory capacity and affective content of stimuli. While the first chapter clearly indicated that high working memory capacity leads to more efficient retrieval, the last two chapters showed the impact of emotion on controlled retrieval depended on the type of information to be retrieved (item vs contextual). The outputs of this dissertation contribute to the developing literature on the efficient use of capacity limited working memory.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Badre, D., Lebrecht, S., Pagliaccio, D., Long, N. M., & Scimeca, J. M. (2014). Ventral Striatum and the Evaluation of Memory Retrieval Strategies. *Journal of Cognitive Neuroscience*, *26*(9), 1928–1948. https://doi.org/10.1162/jocn_a_00596

Badre, D., & Wagner, A. D. (2005). Frontal lobe mechanisms that resolve proactive interference. *Cerebral Cortex*, *15*(12), 2003–2012. https://doi.org/10.1093/Cercor/Bhi075

Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia*, *45*(13), 2883–901. https://doi.org/10.1016/j.neuropsychologia.2007.06.015

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3). https://doi.org/10.1016/j.jml.2012.11.001

Barrett, L. F., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychol Bull*, *130*(4), 553–73. https://doi.org/10.1037/0033-2909.130.4.553

Benjamin, A. S., & Bjork, R. A. (2000). On the relationship between recognition speed and accuracy for words rehearsed via rote versus elaborative rehearsal. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *26*(3), 638–648.

Bisby, J. A., & Burgess, N. (2014). Negative affect impairs associative memory but not item memory. *Learn Mem*, *21*(1), 21–7. https://doi.org/10.1101/lm.032409.113

Broadway, J. M., Redick, T. S., & Engle, R. W. (2010). Individual differences in working memory capacity: Control is (in) the goal. In R. Hassin, K. Ochsner, & Y.Trope (Eds.). In *Self Control in society, mind, and brain.* (pp. 163–174). Oxford: Oxford University Press.

Buchanan, T. W. (2007). Retrieval of emotional memories. *Psychol Bull*, *133*(5), 761–79. https://doi.org/10.1037/0033-2909.133.5.761

Buchanan, T. W., Etzel, J. A., Adolphs, R., & Tranel, D. (2006). The influence of autonomic arousal and semantic relatedness on memory for emotional words. *Int J Psychophysiol*, *61*(1), 26–33. https://doi.org/10.1016/j.ijpsycho.2005.10.022

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., … Riddell, A. (in press). Stan: A probabilistic programming language. Journal of Statistical Software (in press). *Journal of Statistical Software*.

Carrasco, M., McElree, B., Denisova, K., & Giordano, A. M. (2003). Speed of visual processing increases with eccentricity. *Nature Neuroscience*, *6*(7), 699–700. https://doi.org/10.1038/nn1079

Chiu, Y. C., Dolcos, F., Gonsalves, B. D., & Cohen, N. J. (2013). On opposing effects of emotion on contextual or relational memory. *Front Psychol*, *4*, 103. https://doi.org/10.3389/fpsyg.2013.00103

Chow, M., & Conway, A. R. A. (2015). The scope and control of attention: Sources of variance in working memory capacity. *Memory & Cognition*, *43*(3), 325–339. https://doi.org/10.3758/s13421-014-0496-9

Cokely, E. T., Kelley, C. M., & Gilchrist, A. L. (2006). Sources of individual differences in working memory: Contributions of strategy to capacity. *Psychonomic Bulletin & Review*, *13*(6), 991–997. https://doi.org/10.3758/BF03213914

Conway, A. R. A., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: the importance of working memory capacity. *Psychon Bull Rev*, *8*(2), 331–5.

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, *30*(2), 163–183. https://doi.org/10.1016/S0160-2896(01)00096-4

Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends Cogn Sci*, *7*(12), 547–52.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450–466.

D’Argembeau, A., & Van der Linden, M. (2005). Influence of emotion on memory for temporal information. *Emotion*, *5*(4), 503–7. https://doi.org/10.1037/1528-3542.5.4.503

Davachi, L., Mitchell, J. P., & Wagner, A. D. (2003). Multiple routes to memory: distinct medial temporal lobe processes build item and source memories. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(4), 2157–2162. https://doi.org/10.1073/pnas.0337195100

Denkova, E., Wong, G., Dolcos, S., Sung, K., Wang, L., Coupland, N., & Dolcos, F. (2010). The impact of anxiety-inducing distraction on cognitive performance:

a combined brain imaging and personality investigation. *PLoS One*, *5*(11), e14150. https://doi.org/10.1371/journal.pone.0014150

D'Esposito, M., Postle, B. R., Jonides, J., & Smith, E. E. (1999). The neural substrate and temporal dynamics of interference effects in working memory as revealed by event-related fMRI. *Proceedings of the National Academy of Sciences of the USA*, *96*, 7514–7519.

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2007). Imaging recollection and familiarity in the medial temporal lobe: a three-component model. *Trends in Cognitive Sciences*, *11*(9), 379–386. https://doi.org/10.1016/j.tics.2007.08.001

Dobbins, I. G., Rice, H. J., Wagner, A. D., & Schacter, D. L. (2003). Memory orientation and success: separable neurocognitive components underlying episodic recognition. *Neuropsychologia*, *41*(3), 318–333.

Doerksen, S., & Shimamura, A. P. (2001). Source Memory Enhancement for Emotional Words. *Emotion*, *1*(1), 5–11. https://doi.org/10.1037/1528-3542.1.1.5

Dolcos, F., & Denkova, E. (2014). Current Emotion Research in Cognitive Neuroscience: Linking Enhancing and Impairing Effects of Emotion on Cognition. *Emotion Review*, *6*(4), 362–375. https://doi.org/10.1177/1754073914536449

Dolcos, F., Denkova, E., & Dolcos, S. (2012). Neural Correlates of Emotional Memories: A Review of Evidence from Brain Imaging Studies. *Psychologia*, *55*(2), 80–111.

Dolcos, F., Diaz-Granados, P., Wang, L. H., & McCarthy, G. (2008). Opposing influences of emotional and non-emotional distracters upon sustained prefrontal cortex activity during a delayed-response working memory task.

Neuropsychologia, *46*(1), 326–335. https://doi.org/10.1016/J.Neuropsychologia.2007.07.010

Dolcos, F., Iordan, A. D., & Dolcos, S. (2011). Neural correlates of emotion-cognition interactions: A review of evidence from brain imaging investigations. *J Cogn Psychol (Hove)*, *23*(6), 669–694. https://doi.org/10.1080/20445911.2011.594433

Dolcos, F., LaBar, K. S., & Cabeza, R. (2004). Dissociable effects of arousal and valence on prefrontal activity indexing emotional evaluation and subsequent memory: an event-related fMRI study. *NeuroImage*, *23*(1), 64–74. https://doi.org/10.1016/j.neuroimage.2004.05.015

Dolcos, F., & McCarthy, G. (2006). Brain systems mediating cognitive interference by emotional distraction. *J Neurosci*, *26*(7), 2072–9. https://doi.org/10.1523/JNEUROSCI.5042-05.2006

Dosher, B. A. (1981). The effect of delay and interference: A speedaccuracy study. *Cognitive Psychology*, *13*, 551–582.

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience*, *30*, 123–152. https://doi.org/10.1146/annurev.neuro.30.051606.094328

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*(1), 19–23. https://doi.org/10.1111/1467-8721.00160

Engle, R. W. (2010). Role of Working-Memory Capacity in Cognitive Control. *Current Anthropology*, *S1*.

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In *The psychology of learning and motivation* (Ross B, Vol. 44, pp. 145–199). New York: Elsevier.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *J Exp Psychol Gen*, *128*(3), 309–31.

Fenker, D. B., Schott, B. H., Richardson-Klavehn, A., Heinze, H. J., & Duzel, E. (2005). Recapitulating emotional context: activity of amygdala, hippocampus and fusiform cortex during recollection and familiarity. *European Journal of Neuroscience*, *21*(7), 1993–1999. https://doi.org/10.1111/J.1460-9568.2005.04033.X

Feredoes, E., & Postle, B. R. (2010). Prefrontal control of familiarity and recollection in working memory. *J Cogn Neurosci*, *22*(2), 323–30. https://doi.org/10.1162/jocn.2009.21252

Feredoes, E., Tononi, G., & Postle, B. R. (2006). Direct evidence for a prefrontal contribution to the control of proactive interference in verbal working memory. *Proc Natl Acad Sci U S A*, *103*(51), 19530–4. https://doi.org/10.1073/pnas.0604509103

Foraker, S., & McElree, B. (2007). The role of prominence in pronoun resolution: Active versus passive representations. *Journal of Memory and Language*, *56*(3), 357–383. https://doi.org/10.1016/j.jml.2006.07.004

Gardiner, J. M. ; Craik, F. I. M.; Birstwistle, J. (1972). Retrieval cues and release from proactive inhibition. *Journal of Verbal Learning and Verbal Behavior*, *11*, 778–783.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis*. Retrieved from http://public.eblib.com/choice/publicfullrecord.aspx?p=1438153

Gronlund, S. D., Edwards, M. B., & Ohrt, D. D. (1997). Comparison of the retrieval of item versus spatial position information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(5), 1261–1274. https://doi.org/10.1037/0278-7393.23.5.1261

Hacker, M. J. (1980). Speed and accuracy of recency judgments for events in short-term memory. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(6), 651–675. https://doi.org/10.1037/0278-7393.6.6.651

Hamann, S. (2001). Cognitive and neural mechanisms of emotional memory. *Trends in Cognitive Sciences*, *5*, 394–400.

Hintzman, D. L., & Curran, T. (1994). Retrieval Dynamics of Recognition and Frequency Judgments - Evidence for Separate Processes of Familiarity and Recall. *Journal of Memory and Language*, *33*(1), 1–18. https://doi.org/10.1006/Jmla.1994.1001

Hockley, W. E. (1984). Analysis of response time distributions in the study of cognitive processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(4), 598–615. https://doi.org/10.1037/0278-7393.10.4.598

Howard, M. W., & Kahana, M. J. (2002). A Distributed Representation of Temporal Context. *J. Math. Psychol.*, *46*(3), 269–299. https://doi.org/10.1006/jmps.2001.1388

Ilkowska, M., & Engle, R. W. (2010). Trait and State Differences in Working Memory Capacity. In *Handbook of Individual Differences in Cognition:*

Attention, Memory,andExecutive Control, The Springer Series on Human Exceptionality (pp. 295–320). Springer NY.

Iordan, A. D., Dolcos, S., & Dolcos, F. (2013). Neural signatures of the response to emotional distraction: a review of evidence from brain imaging investigations. *Front Hum Neurosci*, *7*, 200. https://doi.org/10.3389/fnhum.2013.00200

Jonides, J., & Nee, D. E. (2006). Brain mechanisms of proactive interference in working memory. *Neuroscience*, *139*(1), 181–93. https://doi.org/10.1016/j.neuroscience.2005.06.042

Jonides, J., Smith, E. E., Marshuetz, C., Koeppe, R. A., & Reuter-Lorenz, P. A. (1998). Inhibition in verbal working memory revealed by brain activation. *Proc Natl Acad Sci U S A*, *95*(14), 8410–3.

Jonides, J., Thompson-Schill, S., D'Esposito, M., Smith, E. E., Marshuetz, C., Kan, I. P., … Swick, D. (2002). Effects of frontal lobe damage on interference effects in working memory. *Cogn Affect Behav Neurosci*, *2*(2), 109–120.

Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *J Exp Psychol Gen*, *130*(2), 169–83.

Kane, M. J., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: limits on long-term memory retrieval. *J Exp Psychol Learn Mem Cogn*, *26*(2), 336–58.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: an individual-differences perspective. *Psychon Bull Rev*, *9*(4), 637–71.

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to Stroop interference. *J Exp Psychol Gen*, *132*(1), 47–70.

Kane, M. J., & McVay, J. C. (2012). Why Does Working Memory Capacity Predict Variation in Reading Comprehension? On the Influence of Mind Wandering and Executive Attention. *Journal of Experimental Psychology: General*, *141*(2), 302–320.

Kensinger, E. A. (2009). Remembering the Details: Effects of Emotion. *Emot Rev*, *1*(2), 99–113. https://doi.org/10.1177/1754073908100432

Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition*, *31*(8), 1169–1180.

Kensinger, E. A., Garoff-Eaton, R. J., & Schacter, D. L. (2006a). Effects of emotion on memory specificity: Memory trade-offs elicited by negative visually arousing stimuli. *Journal of Memory and Language*, *56*(4), 575–591. https://doi.org/10.1016/J.Jml.2006.05.004

Kensinger, E. A., Garoff-Eaton, R. J., & Schacter, D. L. (2006b). Memory for specific visual details can be enhanced by negative arousing content. *Journal of Memory and Language*, *54*(1), 99–112. https://doi.org/10.1016/j.jml.2005.05.005

Kensinger, E. A., & Schacter, D. L. (2008). Memory and emotion, In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.). In *Handbook of emotions* (3rd ed., pp. 601–617). New York: Guilford.

Keppel, G., & Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning and Verbal Behavior*, *1*, 153–161.

Klauer, K. C. (2010). Hierarchical Multinomial Processing Tree Models: A Latent-Trait Approach. *Psychometrika*, *75*(1), 70–98. https://doi.org/10.1007/s11336-009-9141-0

Klein, K. A., Criss, A. H., & Shiffrin, R. M. (2004). Recency judgments and list context. In *Proceedings of the 26th annual meeting of the Cognitive Science Society* (p. 1578). Mahwah, NJ, US.

Klein, K. A., Shiffrin, R. M., & Criss, A. H. (2007). Putting Context in Context. In *The foundations of remembering: Essays in honor of Henry L. Roediger, III* (pp. 171–189). New York, NY, US: Psychology Press.

Knight, M., & Mather, M. (2009). Reconciling findings of emotion-induced memory enhancement and impairment of preceding items. *Emotion*, *9*(6), 763–81. https://doi.org/10.1037/a0017281

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, *14*(4), 389–433. https://doi.org/10.1016/S0160-2896(05)80012-1

LaBar, K. S., & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nat Rev Neurosci*, *7*(1), 54–64. https://doi.org/10.1038/nrn1825

Lang, P. J., Bradley, M. M., & Cuthbert, B. . I. (2005). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual* (Technical Report A-6.). University of Florida, Gainesville, FL.: University of Florida.

Levens, S. M., & Phelps, E. A. (2008). Emotion processing effects on interference resolution in working memory. *Emotion*, *8*(2), 267–80. https://doi.org/10.1037/1528-3542.8.2.267

Levens, S. M., & Phelps, E. A. (2010). Insula and orbital frontal cortex activity underlying emotion interference resolution in working memory. *J Cogn Neurosci*, *22*(12), 2790–803. https://doi.org/10.1162/jocn.2010.21428

Lewandowski, D., Kurowicka, D., & Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis*, *100*(9), 1989–2001. https://doi.org/10.1016/j.jmva.2009.04.008

Linderholm, T. (2002). Predictive Inference Generation as a Function of Working Memory Capacity and Causal Text Constraints. *Discourse Processes*, *34*(3), 259–280. https://doi.org/10.1207/S15326950DP3403_2

Liu, Y. S., Chan, M., & Caplan, J. B. (2014). Generality of a congruity effect in judgements of relative order. *Memory & Cognition*, *42*(7), 1086–1105. https://doi.org/10.3758/s13421-014-0426-x

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: a user's guide* (2nd ed). Mahwah, N.J: Lawrence Erlbaum Associates.

Martin, A. E., & McElree, B. (2009). Memory operations that support language comprehension: Evidence from verb-phrase ellipsis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(5), 1231–1239. https://doi.org/10.1037/a0016271

Mather, M. (2007). Emotional Arousal and Memory Binding An Object-Based Framework. *Perspectives on Psychological Science*, *2*(1), 33–52. https://doi.org/10.1111/J.1745-6916.2007.00028.X

Mather, M., Mitchell, K. J., Raye, C. L., Novak, D. L., Greene, E. J., & Johnson, M. K. (2006). Emotional arousal can impair feature binding in working memory. *J Cogn Neurosci*, *18*(4), 614–25. https://doi.org/10.1162/jocn.2006.18.4.614

Mather, M., & Nesmith, K. (2008). Arousal-Enhanced Location Memory for Pictures. *J Mem Lang*, *58*(2), 449–464. https://doi.org/10.1016/j.jml.2007.01.004

McElree, B. (2001). Working memory and focal attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(3), 817–835. https://doi.org/10.1037/0278-7393.27.3.817

McElree, B. (2006). Accessing Recent Events. In *Psychology of Learning and Motivation* (Vol. 46, pp. 155–200). Elsevier. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0079742106460059

McElree, B., & Carrasco, M. (1999). The temporal dynamics of visual search: Evidence for parallel processing in feature and conjunction searches. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(6), 1517–1539. https://doi.org/10.1037/0096-1523.25.6.1517

McElree, B., Dolan, P. O., & Jacoby, L. L. (1999). Isolating the contributions of familiarity and source information to item recognition: a time course analysis. *J Exp Psychol Learn Mem Cogn*, *25*(3), 563–82.

McElree, B., & Dosher, B. A. (1989). Serial position and set size in shortterm memory: Time course of recognition. *Journal of Experimental Psychology: General*, *18*, 346–373.

McElree, B., & Dosher, B. A. (1993). Serial Retrieval-Processes in the Recovery of Order Information. *Journal of Experimental Psychology-General*, *122*(3), 291–315.

McElree, B., Foraker, S., & Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language*, *48*(1), 67–91. https://doi.org/10.1016/S0749-596X(02)00515-6

McNamara, D. S., & Scott, J. L. (2001). Working memory capacity and strategy use. *Memory & Cognition*, *29*(1), 10–17. https://doi.org/10.3758/Bf03195736

Mecklinger, A., Weber, K., Gunter, T. C., & Engle, R. W. (2003). Dissociable brain mechanisms for inhibitory control: effects of interference content and working memory capacity. *Brain Res Cogn Brain Res*, *18*(1), 26–38.

Mızrak, E., & Öztekin, I. (2016). Relationship between emotion and forgetting. *Emotion (Washington, D.C.)*, *16*(1), 33–42. https://doi.org/10.1037/emo0000069

Monsell, D. (1978). Recency, immediate recognition and reaction time. *Cognitive Psychology*, *10*, 465–501.

Murray, B. D., & Kensinger, E. A. (2012). The effects of emotion and encoding strategy on associative memory. *Memory & Cognition*, *40*(7), 1056–1069. https://doi.org/10.3758/S13421-012-0215-3

Muter, P. (1979). Response latencies in discriminations of recency. *Journal of Experimental Psychology: Human Learning and Memory*, *5*(2), 160–169. https://doi.org/10.1037/0278-7393.5.2.160

Nairne, J. S. (2002). Remembering over the short term: The Case Against the Standard Model. *Annual Review of Psychology*, *53*, 53–81.

Nashiro, K., & Mather, M. (2011). Effects of emotional arousal on memory binding in normal aging and Alzheimer’s disease. *Am J Psychol*, *124*, 301–312.

Nyhus, E., & Badre, D. (2015). Memory Retrieval and the Functional Organization of Frontal Cortex. In D. R. Addis, M. Barense, & A. Duarte (Eds.), *The Wiley Handbook on the Cognitive Neuroscience of Memory* (pp. 131–149). Chichester, UK: John Wiley & Sons, Ltd. Retrieved from http://doi.wiley.com/10.1002/9781118332634.ch7

Oberauer, K. (2005). Binding and Inhibition in Working Memory: Individual and Age Differences in Short-Term Recognition. *Journal of Experimental Psychology: General*, *134*(3), 368–387. https://doi.org/10.1037/0096-3445.134.3.368

Oberauer, K., Farrell, S., Jarrold, C., & Lewandowsky, S. (2016). What limits working memory capacity? *Psychological Bulletin*, *142*(7), 758–799. https://doi.org/10.1037/bul0000046

Ochsner, K. N. (2000). Are affective events richly recollected or simply familiar? The experience and process of recognizing feelings past. *J Exp Psychol Gen*, *129*(2), 242–61. https://doi.org/10.1037/0096-3445.129.2.242

Onal, I., Ozay, M., Mizrak, E., Oztekin, I., & Vural, F. T. Y. (2016). Modeling the Sequence of Brain Volumes by Local Mesh Models for Brain Decoding. *arXiv:1603.01067 [cs]*. Retrieved from http://arxiv.org/abs/1603.01067

Öztekin, I., & Badre, D. (2011). Distributed patterns of brain activity that lead to forgetting. *Frontiers in Human Neuroscience*, *5*, 1–8.

Oztekin, I., Curtis, C. E., & McElree, B. (2009). The medial temporal lobe and the left inferior prefrontal cortex jointly support interference resolution in verbal working memory. *J Cogn Neurosci*, *21*(10), 1967–79. https://doi.org/10.1162/jocn.2008.21146

Oztekin, I., Davachi, L., & McElree, B. (2010). Are Representations in Working Memory Distinct From Representations in Long-Term Memory?: Neural Evidence in Support of a Single Store. *Psychological Science*, *21*(8), 1123–1133. https://doi.org/10.1177/0956797610376651

Öztekin, I., Gungor, N. Z., & Badre, D. (2012). Impact of aging on the dynamics of memory retrieval: A time-course analysis. *Journal of Memory and Language*, *67*(2), 285–294. https://doi.org/10.1016/J.Jml.2012.05.003

Öztekin, I., & McElree, B. (2007). Proactive interference slows recognition by eliminating fast assessments of familiarity. *Journal of Memory and Language*, *57*(1), 126–149. https://doi.org/10.1016/j.jml.2006.08.011

Öztekin, I., & McElree, B. (2010). Relationship Between Measures of Working Memory Capacity and the Time Course of Short-Term Memory Retrieval and Interference Resolution. *Journal of Experimental Psychology-Learning Memory and Cognition*, *36*(2), 383–397. https://doi.org/10.1037/A0018029

Öztekin, I., McElree, B., Staresina, B. P., & Davachi, L. (2008). Working Memory Retrieval: Contributions of the Left Prefrontal Cortex, the Left Posterior Parietal Cortex, and the Hippocampus. *Journal of Cognitive Neuroscience*, *21*(3), 581–593.

Paller, K. A., & Wagner, A. D. (2002). Observing the transformation of experience into memory. *Trends in Cognitive Sciences*, *6*(2), 93–102.

Phelps, E. A., & Sharot, T. (2008). How (and Why) Emotion Enhances the Subjective Sense of Recollection. *Current Directions in Psychological Science*, *17*(2), 147–152.

Piepho, H.-P. (2004). An Algorithm for a Letter-Based Representation of All-Pairwise Comparisons. *Journal of Computational and Graphical Statistics*, *13*(2), 456–466. https://doi.org/10.1198/1061860043515

Pierce, B. H., & Kensinger, E. A. (2011). Effects of emotion on associative recognition: valence and retention interval matter. *Emotion*, *11*(1), 139–44. https://doi.org/10.1037/a0021287

Pitkänen, A., Pikkarainen, M., Nurminen, N., & Ylinen, A. (2009). Reciprocal connections between the amygdala and the hippocampal formation, perirhinal cortex, and postrhinal cortex in rat. A review. *Annals of the New York Academy of Sciences*, *911*(369), 369–391.

Pourtois, G., Schettino, A., & Vuilleumier, P. (2013). Brain mechanisms for emotional influences on perception and attention: what is magic and what is not. *Biol Psychol*, *92*(3), 492–512. https://doi.org/10.1016/j.biopsycho.2012.02.007

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108.

Ratcliff, R., & McKoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Computation*, *20*(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420

Ratcliff, R., & Rouder, J. N. (1998). Modeling Response Times for Two-Choice Decisions. *Psychological Science*, *9*(5), 347–356. https://doi.org/10.1111/1467-9280.00067

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion Decision Model: Current Issues and History. *Trends in Cognitive Sciences*, *0*(0). https://doi.org/10.1016/j.tics.2016.01.007

Ratcliff, R., & Starns, J. J. (2013). Modeling Confidence Judgments, Response Times, and Multiple Choices in Decision Making: Recognition Memory and Motion Discrimination. *Psychological Review*, *120*(3), 697–719. https://doi.org/10.1037/a0033152

Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). Measuring Working Memory Capacity With Automated Complex Span Tasks. *European Journal of Psychological Assessment*, *28*(3), 164–171. https://doi.org/10.1027/15-5759/A000123

Redick, T. S., Calvo, A., Gay, C. E., & Engle, R. W. (2011). Working memory capacity and go/no-go task performance: selective effects of updating, maintenance, and inhibition. *J Exp Psychol Learn Mem Cogn*, *37*(2), 308–24. https://doi.org/10.1037/a0022216

Redick, T. S., & Engle, R. W. (2006). Working memory capacity and attention network test performance. *Applied Cognitive Psychology*, *20*(5), 713–721. https://doi.org/10.1002/acp.1224

Reed, A. V. (1973). Speed-accuracy trade-off in recognition memory. *Science*, *181*(4099), 574–6. https://doi.org/10.1126/science.181.4099.574

Rimmele, U., Davachi, L., Petrov, R., Dougal, S., & Phelps, E. A. (2011). Emotion enhances the subjective feeling of remembering, despite lower accuracy for contextual details. *Emotion*, *11*(3), 553–62. https://doi.org/10.1037/a0024246

Ritchey, M., Dolcos, F., & Cabeza, R. (2008). Role of amygdala connectivity in the persistence of emotional memories over time: an event-related FMRI investigation. *Cereb Cortex*, *18*(11), 2494–504. https://doi.org/10.1093/cercor/bhm262

Schelble, J. L., Therriault, D. J., & Miller, M. D. (2011). Classifying retrieval strategies as a function of working memory. *Memory & Cognition*, *40*(2), 218–230. https://doi.org/10.3758/s13421-011-0149-1

Scimeca, J. M., Katzman, P. L., & Badre, D. (2016). Striatal prediction errors support dynamic control of declarative memory decisions. *Nature Communications*, *7*. https://doi.org/10.1038/ncomms13061

Sharot, T., & Yonelinas, A. P. (2008). Differential time-dependent effects of emotion on recollective experience and memory for contextual information. *Cognition*, *106*(1), 538–547. https://doi.org/10.1016/j.cognition.2007.03.002

Singmann, H., Bolker, B., & Westfall, J. (2015). *afex: Analysis of Factorial Experiments*.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: applications to dementia and amnesia. *J Exp Psychol Gen*, *117*(1), 34–50.

Spillers, G. J., & Unsworth, N. (2011). Variation in working memory capacity and temporal–contextual retrieval from episodic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(6), 1532–1539. https://doi.org/10.1037/a0024852

Staresina, B. P., & Davachi, L. (2006). Differential encoding mechanisms for subsequent associative recognition and free recall. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *26*(36), 9162–9172. https://doi.org/10.1523/JNEUROSCI.2877-06.2006

Starns, J. J., Ratcliff, R., & McKoon, G. (2012). Evaluating the unequal-variance and dual-process explanations of zROC slopes with response time data and the diffusion model. *Cognitive Psychology*, *64*(1-2), 1–34. https://doi.org/10.1016/j.cogpsych.2011.10.002

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., … Joliot, M. (2002). Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain. *NeuroImage*, *15*(1), 273–289. https://doi.org/10.1006/nimg.2001.0978

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2013). Working memory capacity and retrieval from long-term memory: the role of controlled search. *Memory & Cognition*, *41*(2), 242–254. https://doi.org/10.3758/s13421-012-0261-x

Unsworth, N., & Engle, R. W. (2006). A temporal–contextual retrieval account of complex span: An analysis of errors☆. *Journal of Memory and Language*, *54*(3), 346–362. https://doi.org/10.1016/j.jml.2005.11.004

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychol Rev*, *114*(1), 104–32. https://doi.org/10.1037/0033-295X.114.1.104

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37*(3), 498–505. https://doi.org/10.3758/BF03192720

Unsworth, N., Schrock, J. C., & Engle, R. W. (2004). Working memory capacity and the antisaccade task: individual differences in voluntary saccade control. *J Exp Psychol Learn Mem Cogn*, *30*(6), 1302–21. https://doi.org/10.1037/0278-7393.30.6.1302

Unsworth, N., & Spillers, G. J. (2010). Working memory capacity: Attention control, secondary memory, or both? A direct test of the dual-component model. *Journal of Memory and Language*, *62*(4), 392–406. https://doi.org/10.1016/j.jml.2010.02.001

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2012). Working memory capacity and retrieval limitations from long-term memory: An examination of differences in accessibility. *The Quarterly Journal of Experimental Psychology*, *65*(12), 2397–2410. https://doi.org/10.1080/17470218.2012.690438

Van Overschelde, J. P., Rawson, K. A., & Dunlosky, J. (2004). Category norms: An updated and expanded version of the Battig and Montague (1969) norms.

Journal of Memory and Language, *50*(3), 289–335. https://doi.org/10.1016/j.jml.2003.10.003

von Bohlen und Halbach, O., & Albrecht, D. (2002). Reciprocal connections of the hippocampal area CA1, the lateral nucleus of the amygdala and cortical areas in a combined horizontal slice preparation. *Neurosci Res*, *44*(1), 91–100.

Wabersich, D., & Vandekerckhove, J. (2014). Extending JAGS: a tutorial on adding custom distributions to JAGS (with a diffusion model example). *Behavior Research Methods*, *46*(1), 15–28. https://doi.org/10.3758/s13428-013-0369-3

Watkins, O. C., & Watkins, M. J. (1975). Build-up of proactive inhibition as a cue overload effect. *Journal of Experimental Psychology: Human Learning and Memory*, *104*, 442–452.

White, C. N., & Poldrack, R. A. (2014). Decomposing bias in different types of simple decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(2), 385–398. https://doi.org/10.1037/a0034851

Wickelgren. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, *41*, 67–85.

Wickelgren, W. A., & Corbett, A. T. (1977). Associative interference and retrieval dynamics in yes–no recall and recognition. *Journal of Experimental Psychology: Human Learning and Memory*, *3*, 189–202.

Wickelgren, W. A., Corbett, A. T., & Dosher, B. A. (1980). Priming and Retrieval from Short-Term-Memory - a Speed Accuracy Trade-Off Analysis. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 387–404. https://doi.org/10.1016/S0022-5371(80)90276-5

Wickens, D. D. (1970). Encoding categories of words: An empirical approach to meaning. *Psychological Review*, *77*, 1–15.

Yntema, D. B., & Trask, F. P. (1963). Recall as a search process. *Journal of Verbal Learning and Verbal Behavior*, *1963*(2), 65–74.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*(3), 441–517. https://doi.org/10.1006/Jmla.2002.2864

Yonelinas, A. P., & Ritchey, M. (2015). The slow forgetting of emotional episodic memories: an emotional binding account. *Trends in Cognitive Sciences*, *19*(5), 259–267. https://doi.org/10.1016/j.tics.2015.02.009

Supplement I

Asymptotic Accuracy

For the asymptotic accuracy measures, the *d',* for the last two response deadlines were averaged. This measure reflects the terminal accuracy level reached, indicating the probability of successful retrieval (McElree, 2001; McElree & Dosher, 1989, 1993; Öztekin & McElree, 2007).

Fig. S1.1 illustrates the asymptotic accuracy for the two groups across conditions. A 2 (Group [HS vs. LS]) x 10 (all test probe combinations) mixed analysis of variance (ANOVA) conducted on asymptotic *d*' indicated that there was a main effect of the group: HSs had higher accuracy levels than LSs [$F(1, 21) = 9.90$, $p < .006$, $\eta_p^2 =0.20$]. There was also a main effect of the test probe combinations $F(3.85, 80.76) = 38.70$, $p < .001$, $\eta_p^2 =0.46$]. More importantly, there was a significant group by condition, $F(3.85, 80.76) = 8.90$, $p < .001$, $\eta_p^2 =0.16$]. Post Hoc analyses revealed that there were reliable group differences for all conditions (see below for statistical comparisons), except for the most recent one (SP-L 5), $p = .9$.

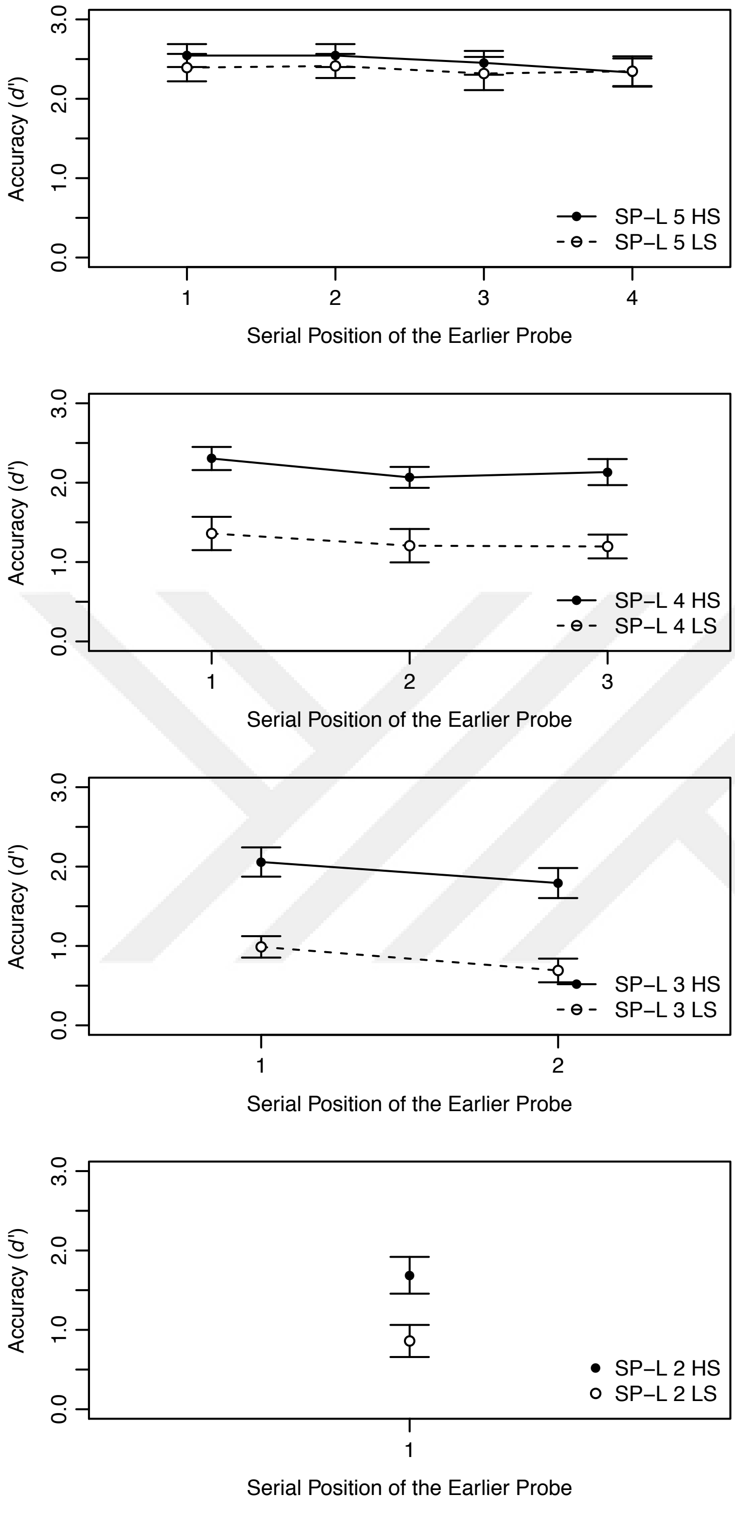

Figure S1.1. Empirical asymptotic accuracy (in *d'* units) measures, derived by averaging over the last two response deadlines, plotted as a function of the study position of the earlier probe for the average high-span (HS) and low-span (LS) groups. Separate lines represent conditions for the more recent probe. Error bars donate 95 % within-subjects confidence intervals. SP-L = Serial position of the later probe.

1.1. Effect of the serial position of the earlier probe

We first investigated the impact of study position of the earlier probe (SP-E), the probe that had been presented earlier in the study list, on accuracy. Discriminating the effect of the earlier and the later probes on memory performance is important, as it is indicative of the strategy applied during temporal order memory retrieval. Consistent with previous investigations of JOR (McElree & Dosher, 1993; Öztekin et al., 2008) we first analyzed the probe conditions holding the later probe constant and varying the earlier probe. A 2 (Group [HS vs. LS]) x 4 (SP-L: 5, SP-E: 1, 2, 3, 4) mixed ANOVA indicated that the serial position of the earlier probe did not have a measurable effect on asymptotic accuracy for SP-L 5 ($p = .22$). Similarly, there was not a reliable main effect of WMC ($p = .87$) or a reliable SP-E by Group interaction ($p = .46$). While SP-E not affecting the recency judgments is consistent with the literature suggesting that order information is recovered with a self-terminating serial scan, the lack of differences between HSs and LSs deserves further explanation. Aforementioned WMC group differences were observed for all positions except for the probes containing the most recently studied item. This finding is consistent with previous research indicating the most recently studied item is maintained in the current focus of attention (FoA) hence does not require an effortful search through memory representations (McElree, 2006; Mızrak & Öztekin, 2016; Öztekin et al., 2012; Öztekin & McElree, 2007, 2010). Accordingly, it is not surprising that WMC effects on asymptotic accuracy were prominent when controlled processing is required, and not present when the memory judgment entailed matching the probe to the contents of focal attention.

While availability of temporal order information did not reliably differ across

HSs and LSs for the most recent item, group differences did emerge for the remaining test probes. For instance, when the SP-L was the 4th item from the study list, A 2 (Group [HS vs. LS]) x 3 (SP-E: 1, 2, 3 compared with SP-L 4) mixed ANOVA analysis showed that WMC had a significant impact on retrieval success [$F(1, 21) = 16.73$, $p < .0006$, $\eta_p^2 = 0.38$], with HSs exhibiting higher accuracy than LSs. We did not observe a significant Group x SP-E interaction ($p = .88$, $\eta_p^2 = 0.001$), or a main effect of SP-E on the asymptotic accuracy ($p = .08$, $\eta_p^2 = 0.03$). Similarly, a 2 (Group [HS vs. LS]) x 2 (test probes containing SP-L 3 [i.e., 31, 32]) revealed that both WMC [$F(1, 21) = 21.64$, $p < .0002$, $\eta_p^2 = 0.48$] and SP-E [$F(1,21) = 13.79$, $p < .002$, $\eta_p^2 = 0.06$] had a significant effect on retrieval success. Accuracy for 3-1 probe combination (M = 1.56) was greater than 3-2 probe combination ($M = 1.27$) indicating an effect of the position of the earlier probe. This effect, however, is not sufficient to construe that the distance between the items impact the recency judgments as it was not observed in the other test probes.

1.2. *Effect of the serial position of the later probe*

After establishing that the SP-E did not modulate empirical accuracy measures, we did a similar analysis by holding SP-E constant across SP-L conditions in order to monitor how memory performance changed with the recency of SP-L. A 2 (Group [HS vs. LS]) x 4 (SP-L: 5, SP-E: 1, 2, 3, 4) mixed ANOVA analysis on the asymptotic accuracy [$F(1.76,36.91) = 39.74$, $p < .0001$; $\eta_p^2 = 0.41$] parameters showed that SP-L had a significant impact on the empirical measures of accuracy. A 2 (Group [HS vs. LS]) x 3 (SP-L: 4 compared with SP-E: 2, 3, 4) mixed ANOVA analysis significant main effect of SP-L on asymptotic accuracy [$F(1.76, 36.97) = 74.90$, $p = .0001$, $\eta_p^2 = 0.51$]. Finally, a 2 (Group [HS vs. LS]) x 2 (SP-L: 3 compared with SP-E: 1, 2) demonstrated that SP-L determined the changes in asymptote [$F(1, 21) = 48.34$, $p < .0001$, $\eta_p^2 = 0.40$].

To further investigate this effect and look at group differences, we assessed asymptotic accuracy for the SP-L, collapsing over the serial position of the earlier probe. A 2 (Group [HS vs. LS]) x 4 (SP-L: 2, 3, 4, 5) mixed ANOVA results showed that accuracy increased as the later item in the test probes was drawn from the more recent items in the study list: [$F(1.86, 39.02) = 37.11$, $p < .0001$; $\eta_p^2 = .43$]. There was also a reliable main effect of group, HS group had higher accuracy than LS group, [$F(1, 21) = 14.23$, $p = .001$, $\eta_p^2 = 0.28$]. In addition, this analysis indicated a significant Group x Later Probe interaction: LS participants were less accurate than HS participants for early positions, but this difference was less prominent as the test probe was more recent, $F(1.86, 39.02) = 7.23$, $p < .001$, $\eta_p^2 = 0.13$.

Fig. S1.2 illustrates that accuracy increased as the SP-L was derived from more recent positions, with a significant linear trend for both HSs ($p < .001$) and LSs ($p < .001$) groups. The data further exhibited a reliable quadratic trend in the rise of accuracy as a function of the recency of the SP-L for the LS group, $t(63) = 4.478$, $p < .001$ which was not observed for HS group, $t(63) = 0.252$, $p = .81$. Namely, LSs performed at similar levels for both SP-L 2 and SP-L 3 and their performance notably increased when SP-L was the 4th item from the study list. We tested the differences between the observed significant linear trends and the linear decrease in accuracy as a function of the later probe was significantly stronger for LSs compared to HSs, $t(63)$

= -3.50, *p* < .001. This suggests that the performance change due to recency of the later probe affected LSs' performance more than HSs.

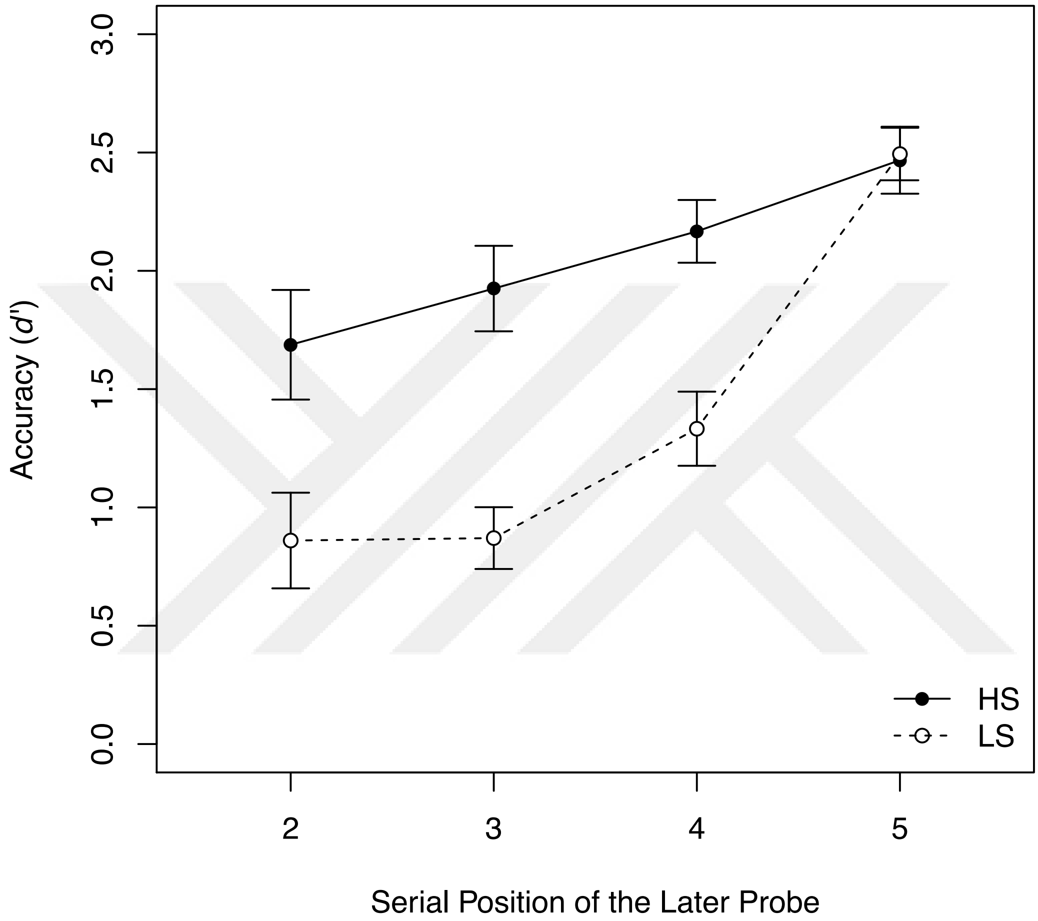

Figure. S1.2. Averaged (over participants within groups) empirical asymptotic accuracy (*d'*) plotted as a function of serial position of the later probe. For each later probe, earlier probes are averaged. Error bars donate 95 % within-subjects confidence intervals.

These asymptotic accuracy differences indicate that LS individuals overall have lower probability of retrieval success in temporal order memory retrieval compared to HS individuals. LSs were specifically impaired in recovering temporal order information when it was effortful. That is to say, this difference in memory performance across the two groups is more prominent for the early members of the study list compared with more recent probes. This is consistent with previous investigations of item recognition that showed the WMC impact was more prominent when the recognition probe was from earlier serial positions (Öztekin & McElree, 2010).

1.3. Primacy effect

Upon observing SP-E modulating the asymptotic accuracy for probe combinations 3-1 and 3-2, one could be concerned that there might be a primacy effect while judging the recency of the items. We did further analysis to assess potential primacy effects in the data. We collapsed SP-L (2,3,4,5) conditions across SP-E (1,2,3,4) as we did for SP-L above and examined whether there is an advantage of having the first item presented in the study list as SP-E. A 2 (Group [HS vs. LS]) x 4 (SP-E: 1, 2, 3, 4) mixed ANOVA showed that SP-E had a significant impact on asymptotic accuracy [$F(1.66, 34.81) = 18.96$, $p < .001$; $\eta_p^2 = 0.22$]. However, when this effect was further investigated by post-hoc tests, we observed that asymptotic accuracy was actually lower or at similar levels for SP-E 1 (Mean; HS = 2.15, LS = 1.47) compared to SP-E 2 (Mean: HS = 2.15, LS = 1.51), SP-E 3 (Mean: HS = 2.30, LS = 1.88), and SP-E 4 (Mean: HS = 2.48, LS = 2.33). Accordingly, we conclude that there was no measurable primacy effects in our data.

Table S1.1A

Asymptote parameter estimates from the 10λ-10β-10δ fit for High Span Group

	Asymptote Parameters									
	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}
Average	1.71	2.16	2.02	2.35	2.45	2.32	2.45	2.42	2.36	2.32
Participants										
1	0.41	2.48	1.37	2.57	2.48	2.14	2.89	2.85	2.85	2.91
2	2.91	2.91	2.70	2.35	2.65	2.91	2.89	2.91	2.87	2.82
3	1.01	1.53	1.76	2.75	2.19	2.68	2.10	2.49	1.87	2.17
4	0.53	2.12	2.15	2.20	2.16	1.88	2.01	1.98	2.14	2.33
5	1.51	2.73	2.53	2.78	2.91	2.90	2.80	2.85	2.61	2.91
6	2.39	1.40	2.91	2.91	2.41	1.98	2.70	2.74	2.76	2.70
7	0.79	2.59	2.44	2.72	2.38	2.91	2.89	2.76	2.80	2.91
8	0.87	0.78	0.60	1.22	0.96	0.95	1.44	1.28	1.30	1.37
9	2.91	2.72	2.60	2.82	2.89	2.68	2.88	2.91	2.91	1.67
10	2.17	1.75	1.64	1.85	1.82	1.76	2.16	2.08	1.95	2.03
11	2.21	2.55	1.50	2.91	2.91	2.91	2.89	2.67	2.91	2.71
12	2.07	2.14	2.03	2.89	2.71	1.80	2.46	2.47	2.89	2.42

Table S1.1B

Asymptote parameter estimates from the 10λ-10β-10δ fit for Low Span Group

	Asymptote Parameters									
	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}
Average	0.85	1.08	0.92	1.60	1.35	1.29	2.43	2.41	2.41	2.30
Participants										
1	1.68	2.07	0.78	2.47	1.97	1.82	2.89	2.72	2.66	1.94
2	0.27	0.70	0.60	0.56	1.20	1.00	1.82	1.58	1.94	1.95
3	0.52	1.11	0.75	2.91	1.26	1.73	2.56	2.46	2.64	2.43
4	1.10	0.99	1.32	1.41	0.97	1.22	2.07	2.07	2.09	2.80
5	0.38	0.87	0.32	0.98	1.50	1.08	1.74	1.64	1.57	1.86
6	0.37	0.46	0.34	0.34	1.22	0.53	2.36	2.66	2.53	2.53
7	1.60	1.01	1.25	1.10	1.05	1.09	2.62	2.67	2.67	2.59
8	0.22	1.25	1.37	1.43	1.61	1.23	2.58	2.62	2.63	2.13
9	1.45	0.64	0.01	1.87	0.13	0.83	2.89	2.85	2.91	2.83
10	2.27	2.46	2.68	2.43	2.68	2.24	2.91	2.91	2.91	2.91
11	0.36	1.73	0.52	1.61	1.61	2.91	2.75	2.87	2.87	2.75

Note. Average parameters are based on data averaged over participants (not the average of individual parameters). Asymptote numbers from 1 to 10 refers to the asymptote measures estimated by the full fit for probe combinations; 2-1, 3-1, 3-2, 4-1, 4-2, 4-3, 5-1, 5-2, 5-3, 5-4 respectively.

Table S1.2A

Rate parameter estimates from the 10λ-10β-10δ fit for High Span Group

	Rate Parameters									
	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}
Average	1.28	1.83	1.26	2.44	0.94	1.20	10.49	6.57	5.98	3.77
Participants										
1	8.52	2.41	11.37	3.38	1.73	1.74	8.91	20.00	20.00	20.00
2	0.92	1.26	2.15	7.21	1.30	1.00	17.47	20.00	20.00	7.42
3	8.67	2.04	1.06	0.53	0.87	3.05	20.00	9.83	20.00	20.00
4	10.59	3.45	3.45	4.53	2.28	10.80	20.00	20.00	2.35	1.58
5	11.43	0.96	1.48	1.82	0.85	1.35	3.21	1.64	3.79	1.53
6	1.27	2.50	0.23	1.44	0.74	0.82	13.00	20.00	19.99	20.00
7	4.42	4.72	5.21	2.68	2.71	2.71	4.34	8.17	3.78	2.52
8	4.58	20.00	5.28	10.56	5.68	4.59	14.86	11.72	5.15	4.82
9	5.21	15.25	1.97	15.03	2.02	2.88	13.33	4.26	1.31	0.76
10	2.83	5.85	20.00	20.00	14.78	3.24	2.68	1.88	10.93	3.40
11	2.22	1.50	17.31	1.05	0.82	0.45	3.33	11.00	2.53	2.51
12	2.22	2.35	9.61	0.82	0.98	4.74	16.59	20.00	11.56	5.24

Note. Average parameters are based on data averaged over participants (not the average of individual parameters).

Table S1.2B

Parameter estimates from the 10λ-10β-10δ fit for Low Span Group

	Rate Parameters									
	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}
Average	5.42	1.72	0.82	1.27	2.11	1.85	10.81	7.42	8.82	9.41
Participants										
1	11.89	0.94	9.46	2.17	2.44	0.51	2.55	3.28	1.45	1.49
2	9.08	12.26	13.47	10.41	1.12	1.70	20.00	20.00	3.92	1.96
3	13.12	1.78	17.00	0.60	11.83	14.47	5.13	20.00	3.72	11.11
4	3.27	0.76	4.32	4.53	13.75	14.70	12.33	11.77	13.61	1.47
5	8.63	10.43	10.61	10.37	0.55	1.81	20.00	12.93	20.00	20.00
6	13.77	19.04	16.97	20.00	3.36	4.71	20.00	4.81	20.00	20.00
7	2.58	2.99	2.48	2.83	1.99	1.57	20.00	14.49	9.71	20.00
8	2.00	6.65	12.02	20.00	20.00	1.56	7.40	6.06	4.97	20.00
9	5.81	9.54	0.10	1.04	11.11	12.84	14.10	20.00	20.00	20.00
10	14.70	0.65	5.05	12.47	5.90	9.33	20.00	5.40	7.29	20.00
11	0.53	0.52	5.97	2.84	4.15	0.29	10.13	4.15	13.17	2.72

Note. Average parameters are based on data averaged over participants (not the average of individual parameters). Rate numbers from 1 to 10 refers to the asymptote measures estimated by the full fit for probe combinations; 2-1, 3-1, 3-2, 4-1, 4-2, 4-3, 5-1, 5-2, 5-3, 5-4 respectively.

Table S1.3A

Intercept parameter estimates from the 10λ-10β-10δ fit for High Span Group

	Intercept Parameters									
	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	δ_8	δ_9	δ_{10}
Average	0.40	0.32	0.33	0.27	0.11	0.28	0.30	0.25	0.27	0.27
Participants										
1	1.92	0.47	0.45	0.28	0.22	0.25	0.21	0.28	0.29	0.30
2	0.31	0.28	0.65	0.30	0.01	0.09	0.27	0.27	0.28	0.24
3	0.94	0.57	0.90	0.17	0.53	0.92	0.36	0.34	0.33	0.37
4	1.00	0.32	0.29	0.36	0.31	0.66	0.31	0.31	0.19	0.01
5	0.99	0.01	0.23	0.25	0.02	0.20	0.24	0.18	0.29	0.32
6	0.52	0.52	0.38	0.40	0.30	0.14	0.27	0.30	0.32	0.31
7	0.65	0.28	0.33	0.30	0.19	0.57	0.23	0.27	0.26	0.24
8	0.66	0.47	0.34	0.19	0.25	0.37	0.32	0.01	0.27	0.34
9	0.27	0.39	0.51	0.37	0.51	0.67	0.37	0.64	0.65	0.51
10	0.64	0.60	0.69	0.39	0.01	0.01	0.01	0.01	0.37	0.48
11	0.59	0.53	1.10	0.33	0.81	0.45	0.15	0.27	0.15	0.21
12	0.62	0.59	0.98	0.16	0.36	0.63	0.33	0.35	0.36	0.32

Note. Average parameters are based on data averaged over participants (not the average of individual parameters).

Table S1.3B

Intercept parameter estimates from the 10λ-10β-10δ fit for Low Span Group

	Intercept Parameters									
	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	δ_8	δ_9	δ_{10}
Average	0.96	0.54	0.50	0.31	0.50	0.56	0.32	0.29	0.31	0.32
Participants										
1	0.89	0.19	1.02	0.39	0.56	0.34	0.26	0.55	0.31	0.21
2	1.00	0.68	0.70	0.32	0.42	0.46	0.33	0.29	0.17	0.19
3	1.03	0.53	0.48	0.17	0.67	0.65	0.22	0.35	0.19	0.33
4	1.32	0.01	1.54	0.86	0.99	0.97	0.27	0.27	0.29	0.24
5	1.15	0.98	0.41	0.58	0.42	0.26	0.34	0.31	0.36	0.40
6	0.97	0.71	0.67	0.01	0.60	0.58	0.38	0.28	0.38	0.37
7	0.95	0.72	1.33	0.54	0.36	0.41	0.36	0.33	0.33	0.35
8	2.80	0.69	1.65	0.70	0.69	0.34	0.33	0.29	0.31	0.36
9	0.75	1.01	2.00	0.24	0.57	0.74	0.36	0.36	0.37	0.37
10	1.63	0.51	1.64	0.96	0.95	0.96	0.36	0.28	0.25	0.35
11	2.77	0.50	0.55	0.38	0.39	0.01	0.07	0.01	0.29	0.01

Note. Average parameters are based on data averaged over participants (not the average of individual parameters). Intercept numbers from 1 to 10 refers to the asymptote measures estimated by the full fit for probe combinations; 2-1, 3-1, 3-2, 4-1, 4-2, 4-3, 5-1, 5-2, 5-3, 5-4 respectively.

Table S1.4

Adjusted- R^2 values for individual fits and average fit from the 10λ-10β-10δ fit for Low Span and High Span groups.

	High R^2	Low R^2
Average	0.964	0.971
Participants		
1	0.912	0.862
2	0.882	0.858
3	0.674	0.870
4	0.887	0.817
5	0.792	0.827
6	0.925	0.883
7	0.922	0.951
8	0.693	0.854
9	0.913	0.937
10	0.805	0.924
11	0.888	0.917
12	0.883	

Note. Average parameters are based on data averaged over participants (not the average of individual parameters).

2. Best fitting SAT models for temporal order memory retrieval

After we established that the earlier probe did not reliably affect temporal order memory performance, we fit the data from both groups with sets of nested and non-nested models that systematically varied the 3 parameters in the exponential function shown in main text Equation 1. Models ranged from a null model consisting of functions fitted with a single asymptote (λ), rate (β), and intercept (δ) across SPs of the later probes (i.e., three parameters in total), to a fully saturated model in which the four later probe conditions were fit with a unique asymptote (λ), rate (β), and intercept (δ) (i.e., 12 parameters in total). Using our knowledge about the variations in the data as a function of the position of the later probe, we then fit the data by varying the parameters across SP-L conditions. We tested the whole range of models starting from the null model described above to the fully saturated model that allocated separate parameters to each SP-L condition. For instance, the 4-1-1 model allocated separate asymptotes to SP-L 2, SP-L 3, SP-L 4, and SP-L 5 conditions, while a common intercept and rate parameter were shared between the conditions (i.e., six parameters in total). We exhaustively explored the full space of possible models by estimating eleven variations for each of the three SAT parameters for a total of 11 x 11 x 11 = 1331 models. These variations were:

1) A common parameter for all, 2) A common parameter for the SP-L 2, SP-L 3 and SP-L 4, and another parameter to SP-L 5 conditions, 3) A common parameter to SP-L 2 and all SP-L 3 conditions, another parameter to all SP-L 4 and all SP-L 5 conditions 4) One parameter to SP-L 2 condition, another parameter for all SP-L 3, SP-L 4 and SP-L 5 conditions, 5) A common parameter to SP-L2 condition and 3-2 probe combination, another parameter to 3-1 probe combination, and all SP-L 4, SP-L 5 conditions 6) One parameter to SP-L 2 condition, one parameter to all SP-L 3 conditions, and a common parameter to all SP-L 4 and SP-L 5 conditions 7) One parameter for SP-L 2 condition and 3-2 probe combination, a common parameter for 3-1 probe combination and all SP-L 4 conditions, and another parameter for all SP-L conditions, 8) One parameter for SP-L 2, a common parameter for SP-L 3 and SP-L 4 conditions, and another parameter for SP-L conditions, 9) A common parameter for SP-L 2 and SP-L 3, one parameter for SP-L 4, and another parameter for SP-L 5 conditions, 10) A common parameter for SP-L 2 condition and 3-2 probe combination, a common parameter for 3-1 probe combination and SP-L 4 conditions, and another parameter for SP-L 5 conditions, 11) One parameter for SP-L 2 condition, one parameter for SP-L 3 conditions, one parameter for SP-L 4 conditions, and another parameter for SP-L 5 conditions. As model selection criteria, weighing model fit against model flexibility we used Adjusted-R^2 (Reed, 1973). To avoid local minima, we ran 50 fitting runs with different random starting values for each individual data

set. In addition to the individual data sets, we also fitted the average data (averaged across participants).

Results showed that there was not a single model that provided a clearly superior account. Instead, there were around 100 models that explained each group's data well. These models fit the average data (averaged across participants) from LS group within the range of .96 - .97 adjusted-R^2, and the mean R^2 of the individual fits from each participant's data were between .86 - .87 for these models. The best models for HS group had R^2 values of .917 - .927 for the average data, and .79 - .80 for the mean R^2 of the individual fits. We examined the shared qualities between these models and identified distinct patterns that improved the adjusted-R^2 values, for asymptote, rate, and intercept parameters (see Fig.s S1.3 and S1.4). Firstly, none of the best models allocated a common parameter to all SP-L conditions, neither for the asymptote parameter nor for the intercept and rate parameters. Additionally, at least two third of the best models allocated a separate parameter for the SP-L 5 condition, which consists of the last item presented in the study list as the later probe. This is consistent with previous research (e.g., McElree & Dosher, 1989, 1993; Mızrak & Öztekin, 2016; Öztekin & McElree, 2007, 2010), suggesting that allocating a separate retrieval speed and/or retrieval accuracy parameter significantly improves the adjusted-R^2 statistics across participants. The following Fig.s depict the R^2 values of the models as a function of the number of parameters.

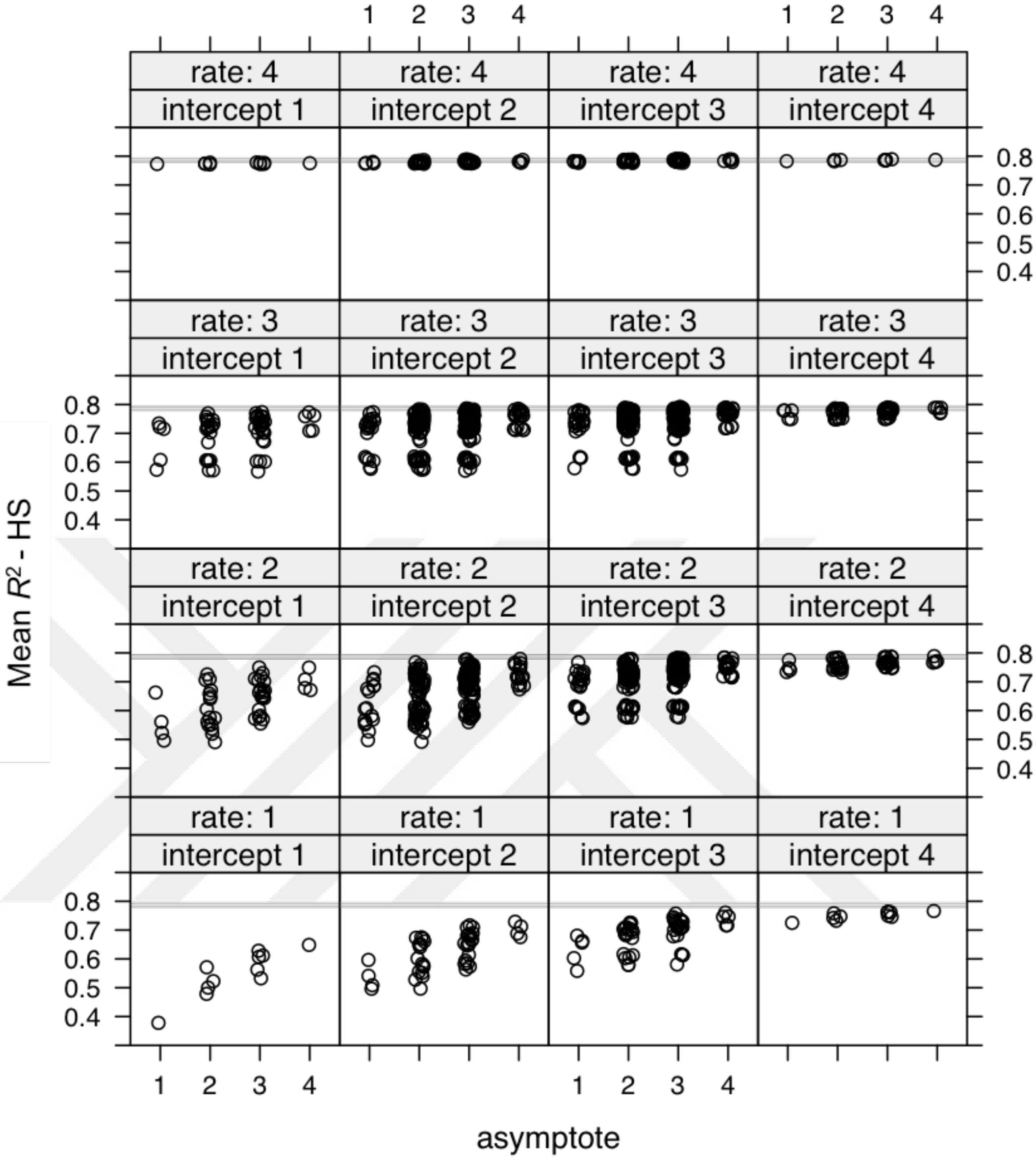

Figure S1.3. R^2 values from all the tested models (on the y-axis), which predicted the empirical data of High-span group by using Equation 1 (main text). Each data point corresponds to the mean R^2 value of the individual fits for a specific model. The models varied each parameter from Equation 1 as a function of the SP-L. The number of asymptote parameters allocated by the model is presented on the x-axis, the number of rate parameters increase across rows from the bottom to the top, and the number of intercept parameter increase across columns from left to right. The two grey lines show the lower and upper threshold for the best fitting models, which were .79 and .80 respectively. As is apparent from the Fig., the R^2 values increased when the models allocated different parameters to different Later Probe conditions. This was more prominent when the retrieval speed parameters (intercept and rate) were varied.

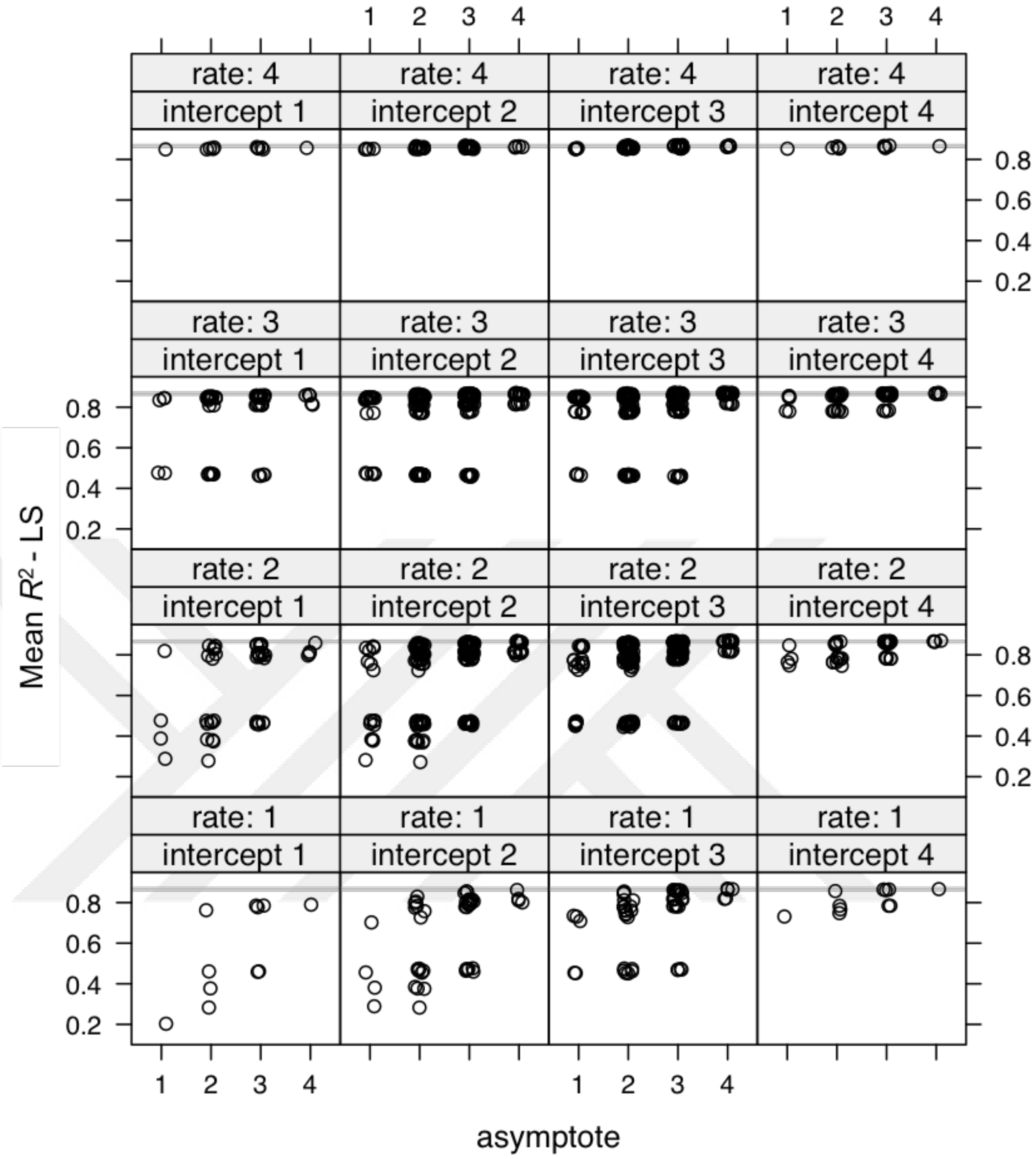

Figure S1.4. R^2 values from all the tested models (on the y-axis), which predicted the empirical data of Low-span group by using the Equation 1. The models varied each parameter from the Equtation 1 as a function of SP-L. Variations in the asymptote parameter (on the x-axis), variations in the rate parameter (increasing across rows from the bottom to the top), and intercept parameter (increasing across columns from left to right). The two grey lines resemble the lower and upper threshold for the best fitting models, which were .86 and .87 respectively. R^2 values increased when the models allocated different parameters to different Later Probe conditions. Similar to the HS group's model fits, this was more prominent when the retrieval speed parameters (intercept and rate) were varied.

Although in our case, there is no single "best" model, we picked representative models amongst the best models described above that fit the data well. We chose the models that had the highest mean R^2 of the individual fits for each group. For HS group, the 3λ-3β-3δ model was the model which represented the empirical data best. This model allocated one asymptote for SP-L 2 condition and 3-2 probe combination (λ_1), and another asymptote for the 3-1 probe combination, and SP-L 4 conditions (λ_2), while allocating a separate parameter to the SP-L 5 conditions (λ_3). According to the best fitting model, three rate parameters were fit to the data of HS group; one for SP-L 2 condition (β_1), one for SP-L 3 and SP-L 4 conditions (β_2), and one for SP-L 5 conditions (β_3). As for the intercept parameter, 3λ-3β-3δ model dictated three intercept parameters for the data; one for SP-L 2 and SP-L 3 conditions (δ_1), one for SP-L 4 conditions (δ_2) and another one for SP-L 5 conditions (δ_3).

For the LS group, 3λ-4β-3δ model allocated three asymptotes; one asymptote for SP-L 2 and 3-2 combinations (λ_1), and another asymptote for the 3-1, and SP-L 4 conditions (λ_2), while allocating a separate parameter to the SP-L 5 conditions (λ_3). The rate parameter also varied as a function of the later probe. Each SP-L condition was allocated a separate rate parameter; SP-L 2(β_1),SP-L 3(β_2),SP-L 4 (β_3),SP-L 5(β_4). Finally, the model assigned 3 different intercept parameters; one for SP-L 2 condition, and 3-2 probe combination (δ_1), one for SP-L 4 conditions and 3-1 probe combinations (δ_2), and one for SP-L 5 conditions (δ_3). The parameter values derived from the models are presented in Table 1.4A for HSs and Table 1.4B for LSs. It is apparent from Fig.s S1.1 and S1.2 that adding extra parameters to these models (i.e., the 4-4-4 models), did not yield further improvements for any of the participants.

One of the criteria for the model to describe the data well is its ability to explain the systematic deviations that are observed in the data. When rates or intercepts were varied among later probes, it both improved the quality of the fits and yielded a set of consistently ordered parameter estimates across the fits of the individual participants' data (see next section for details). Fig. S1.5 illustrates that the models provide a good fit with no apparent systematic deviations.

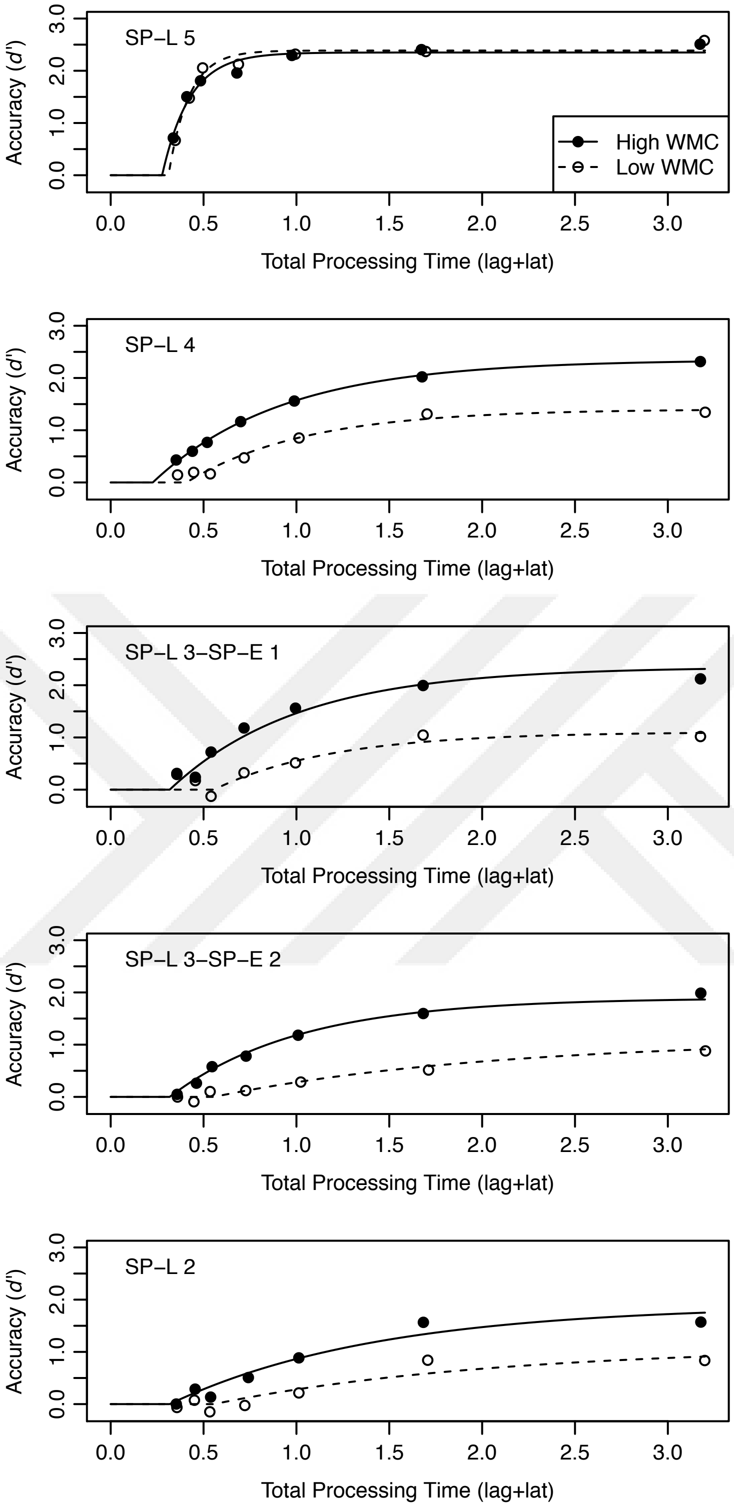

Figure S1.5. Accuracy (in *d'* units) plotted for test probe combinations as a function of the later probe against total processing time (in seconds) for the average high-span (HS) and low-span (LS) groups. For each later probe except the SP-L 3 consisting of the 3-2 and 3-1 probe combinations, SP-E did not have an effect on the SAT function parameters. The symbols indicate empirical data points, and the smooth curves indicate the model fits derived from Equation 1.

Table S1.5A

Parameter estimates from the 3λ-3β-3δ fit for High Span Group

	Parameters									
	δ_1	δ_2	δ_3	β_1	β_2	β_3	δ_1	δ_2	δ_3	R^2
Average	1.873	2.302	2.369	0.92	1.50	6.42	0.322	0.232	0.274	0.927
Participants										
1	1.692	2.397	2.863	0.10	2.20	20	0.426	0.251	0.287	0.900
2	2.732	2.841	2.826	1.03	1.22	19.32	0.317	0.038	0.274	0.879
3	1.640	2.502	2.131	0.62	0.71	20	0.569	0.337	0.363	0.731
4	2.166	2.073	2.105	0.10	3.49	3.66	0.304	0.397	0.182	0.840
5	2.579	2.844	2.788	0.33	1.13	2.15	0.104	0.123	0.241	0.805
6	2.100	2.910	2.708	1.57	0.50	20	0.523	0.081	0.308	0.910
7	2.491	2.637	2.849	0.18	3.15	3.84	0.253	0.370	0.237	0.859
8	0.695	0.936	1.336	2.23	16.94	6.70	0.473	0.342	0.267	0.572
9	2.628	2.910	2.656	6.15	1.73	1.15	0.263	0.298	0.244	0.601
10	1.818	1.735	2.003	4.88	11.91	4.30	0.686	0.242	0.317	0.655
11	2.053	2.892	2.700	2.06	0.79	4.70	0.520	0.521	0.221	0.869
12	2.140	2.351	2.564	1.99	1.42	10.22	0.597	0.369	0.337	0.879

Note. Average parameters are based on data averaged over participants (not the average of individual parameters).

Table S1.5B

Parameter estimates from the 3λ-4β-3δ fit for Low Span Group.

	Parameters										
	δ_1	δ_2	δ_3	β_1	β_2	β_3	β_4	δ_1	δ_2	δ_3	R^2
Average	0.965	1.377	2.387	1.53	0.81	1.60	9.05	0.759	0.411	0.311	0.971
Participants											
1	1.427	2.041	2.617	20	1.09	1.21	1.56	0.911	0.325	0.258	0.760
2	0.620	0.840	1.732	0.25	6.31	3.81	8.21	0.701	0.627	0.273	0.834
3	1.056	1.692	2.481	0.13	0.72	3.50	8.81	0.154	0.532	0.302	0.847
4	1.406	1.332	2.104	0.84	0.69	1.23	9.14	0.710	0.594	0.289	0.833
5	0.491	0.998	1.690	0.40	1.01	3.02	20	0.348	0.487	0.362	0.839
6	0.628	0.657	2.480	0.39	0.44	1.60	20	0.369	0.067	0.376	0.899
7	1.483	1.098	2.632	3.70	1.05	2.18	14.92	0.969	0.481	0.343	0.954
8	1.360	1.326	2.448	0.10	5.73	14.87	8.43	1.586	0.684	0.328	0.866
9	1.385	0.859	2.873	20	0.10	2.60	20	0.947	0.452	0.371	0.905
10	2.594	2.408	2.901	2.65	1.24	5.24	20	1.378	0.917	0.359	0.939
11	0.775	1.701	2.804	0.12	0.38	1.39	4.36	0.010	0.224	0.010	0.907

Note. Average parameters are based on data averaged over participants (not the average of individual parameters).

3. *Evidence for Serial Scan*

The backwards serial scan applied during the retrieval of temporal order information exhibited itself in both retrieval speed and retrieval accuracy parameters for both groups. We examined the effects of this serial scan from the parameter estimates of best fitting models for each individual participant. We found significant linear trends as a function of the serial position of the later probe. Specifically, linear trends indicated that both groups applied a backwards serial scan mechanism. Namely, both HS and LS groups performed better when the later probe of the test probes was drawn from more recent positions in the study list. There was a significant linear increase in the asymptote parameter (HS [$t(33) = 5.56$, $p < .001$], LS [$t(30)=8.74$, $p < .001$]), showing that the probability of retrieval increased from SP-L 2 combinations to SP-L 5 combinations. The speed of evidence accumulation, measured by the rate parameter, also linearly increased as a function of the SP-L (HS [$t(33) = 3.40$, $p < .01$], LS [$t(30)=3.23$, $p < .01$]). Participants reached the temporal order information for the test probe combinations which consisted more recent probes as the SP-L, shown in a linearly decreasing pattern in the intercept parameter (HS [$t(33) = -4.10$, $p < .001$], LS [$t(30)= - 4.88$, $p < .001$]).

Additionally, SP-L had significant main effects on all of the parameter estimates and for both groups revealed by a within group ANOVA analysis (asymptote: HS[$F(1.86, 20.42) = 10.94$, $p < .001$; $\eta_p^2 = 0.08$], LS[$F(3,30) = 36.16$, $p < .001$; $\eta_p^2 = 0.48$], rate: HS[$F(3,33) = 4.81$, $p < .01$, $\eta_p^2 = 0.24$], $F(2.06,20.61) = 6.91$, $p < .01$, $\eta_p^2 = 0.36$], intercept: HS[$F(3,33) = 6.85$, $p < .01$, $\eta_p^2 = 0.21$], LS[$F(3,30) = 8.06$, $p < .001$; $\eta_p^2 = 0.22$]).

In summary, the data from both groups could be best accounted by models that allocated faster speed of retrieval measures (intercept and rate parameters) and higher levels of asymptote measures for the test probes that contained more recently studied items as the later probe, indicative of a backwards serial memory search strategy.

Supplement II

Within List Dynamics

Initially, we evaluated whether SAT functions for neutral and emotional trials exhibited the same patterns observed in previous studies (McElree, 2006). To do so, SAT functions for the 3 SPs were fit with sets of nested models that systematically varied the 3 parameters in Equation 1. Models ranged from a null model consisting of functions fitted with a single asymptote (λ), rate (β), and intercept (δ) across SPs, to a fully saturated model in which each SP function was fit with a unique asymptote (λ), rate (β), and intercept (δ). The null model that allocated a single asymptote (λ), rate (β), and intercept (δ) was significantly improved by a 3λ-1β-1δ model that allocated three unique asymptotes for the three serial positions (*Adjusted-*R^2 for the average data increased from .909 to .930 and from .914 to .937 for emotional and neutral trials respectively), with a reliable improvement in the *Adjusted-*R^2 statistics across our participants [$t(15) = 2.57, p < .022$ for emotional trials; $t(15) = 2.58, p < .021$ for neutral trials]. In addition, consistent with previous research (e.g., McElree & Dosher, 1989; Öztekin & McElree, 2007; 2010) allocating a unique rate parameter for the most recent serial position further improved the model (*Adjusted-*R^2 for the average data increased from .930 to .951 and from .937 to .958 for emotional and neutral trials respectively), yielding a further significant improvement on the *Adjusted-*R^2 statistics across participants [$t(15 = 3.39, p < .004$ for emotional trials; $t(15) = 3.41, p < .004$ for neutral trials].

Accordingly, among these models, 3λ–2β–1δ and 3λ–1β–2δ provided the best fit of the empirical data. Similar patterns were observed for both stimuli types when dynamics of retrieval were assessed as a function of SP. Specifically, The 3λ–1β–2δ model allocated a common rate to all serial positions, one intercept for SPs

1–2, and another intercept for SP 3. The intercept for the most recently studied probe was significantly faster compared to the intercept for SPs 1-2, [$t(15) = 5.18$, $p < .01$, $d = 1.543$ for emotional trials, and $t(15) = 7.59$, $p < .01$, $d = 1.115$ for neutral trials]. 3*l*-2*b*-1*d* model allocated a separate asymptote to each serial position, one rate for SPs1–2, another rate for SP 3 (the most recently studied item), and a common intercept for all the three serial positions. The rate for the last SP (SP 3) was faster compared to SPs 1 and 2 for both emotional [$t(15) = -6.38$, $p < .01$, $d = 1.775$] and neutral trials [$t(15) = -5.58$, $p < .01$, $d = 1.846$]. In addition, the intercept parameter was slower for emotional compared to neutral trials, $t\ (15) = -2.31$, $p < .036$, $d = 0.580$.This finding is further evaluated in the Results section under; Within and Composite List Dynamics. Parameter estimates for each serial position across participants as well as the average data for E and N trials are reported in Tables S2.1A, S2.1B and S2.2A, S2.2B.

In summary, the data could be best accounted by two models that allocated faster speed of retrieval (by allocating either two intercepts or two rates) for the most recently studied image across both emotional and neutral study material. These results are consistent with previous research that implicates the most recent item benefits from a privileged state in the focus of attention (McElree & Dosher, 1993; Öztekin et al., 2012; Öztekin & McElree, 2007; Wickelgren et al., 1980) and extends this phenomenon to emotional study material.

Table S1.1A

Parameter estimates from the 3λ-2β-1δ serial position fits for emotional trials

	Parameters						
	λ_1	λ_2	λ_3	β_1	β_2	δ	R^2
Average	3.67	3.98	3.98	3.52	6.05	.306	0.95
Participants							
1	3.58	4.11	4.1	5.61	9.05	.359	.9
2	3.97	4.11	4.	3.47	6.18	.378	.942
3	4.17	4.15	4.05	6.07	13.73	.360	.949
4	3.90	4.27	4.08	5.56	11.93	.123	.773
5	3.47	3.70	3.82	13.98	15	.510	.94
6	3.60	3.73	3.92	2.67	2.46	.373	.966
7	3.79	4.25	4.14	4.46	5.41	.395	.959
8	3.31	3.73	3.84	4.97	8.29	.468	.916
9	3.86	3.66	4.	6.82	14.92	.384	.939
10	3.25	4.05	3.91	2.91	6.97	.366	.921
11	3.53	4.00	4.27	5.69	15	.334	.885
12	4.19	4.00	3.97	4.11	10.85	.284	.946
13	2.90	3.63	3.91	6.2	15.4	.366	.938
14	3.41	3.77	3.80	3.74	7.40	.288	.739
15	3.37	3.78	3.92	4.44	9.67	.357	.84
16	3.90	4.15	4.17	5.32	15	.354	.95

Note. Average parameters are based on data averaged over participants (not the average of individual parameters).

Table S2.1B

Parameter estimates from the 3λ-2β-1δ serial position fits for neutral trials

	Parameters						
	λ_1	λ_2	λ_3	β_1	β_2	δ	R^2
Average	3.70	3.98	4.02	3.69	6.03	.29	.957
Participants							
1	3.98	4.18	4.23	6.79	15	.346	.956
2	3.40	3.97	3.89	3.54	6.59	.368	.899
3	3.88	4.35	4.02	5.77	12.14	.299	.941
4	4.13	3.76	4.06	4.68	10.2	.087	.670
5	3.38	3.65	3.95	8.35	15.00	.461	.917
6	3.50	3.86	3.63	4.02	5.54	.406	.932
7	3.60	4.13	3.94	3.93	5.60	.371	.923
8	3.11	3.61	4.19	7.27	6.04	.406	.88
9	4.16	3.96	4.19	5.07	7.59	.340	.902
10	3.34	3.93	4.03	2.72	4.27	.313	.924
11	3.33	3.92	4.11	8.26	14.98	.291	.88
12	3.91	4.13	4.13	6.17	9.98	.265	.971
13	3.53	3.95	4.06	4.19	8.29	.324	.947
14	3.98	3.83	3.80	3.28	12.70	.376	.890
15	3.37	3.65	3.85	3.63	5.02	.305	.842
16	4	4.5	4.26	5.84	15	.354	.901

Note. Average parameters are based on data averaged over participants (not the average of individual parameters).

Table S2.2A

Parameter estimates from the 3λ-1β-2δ serial position fits for emotional trials

	Parameters						
	λ_1	λ_2	λ_3	β_1	δ_1	δ_2	R^2
Average	3.61	3.91	3.99	5.37	.37	.285	.976
Participants							
1	3.56	4.09	4.13	6.76	.38	.328	.914
2	3.9	4.05	4.11	4.23	.4	.3415	.947
3	4.15	4.12	4.12	7.25	.38	.295	.976
4	3.86	4.23	4.12	7.04	.16	.01	.763
5	3.48	3.71	3.79	15	.52	.503	.947
6	3.63	3.75	3.87	2.58	.373	.371	.965
7	3.88	4.34	4.09	3.77	.38	.297	.956
8	3.27	3.72	4.05	3.93	.4	.401	.907
9	3.86	3.66	4.13	6.59	.38	.326	.931
10	3.15	3.94	3.91	5.35	.47	.317	.949
11	3.42	3.84	4.25	10.4	.37	.298	.926
12	4.08	3.9	4.08	6.9	.35	.251	.954
13	2.83	3.53	4.06	7.78	.37	.328	.907
14	3.34	3.69	3.95	4.46	.3	.238	.72
15	3.3	3.69	3.9	8.28	.42	.331	.902
16	3.82	4.07	4.23	7.5	.38	.287	.953

Note. Average parameters are based on data averaged over participants (not the average of individual parameters).

Table S2.2B

Parameter estimates from the 3λ-1β-2δ serial position fits for neutral trials

	Parameters						
	λ_1	λ_2	λ_3	β_1	δ_1	δ_2	R^2
Average	3.63	3.91	4	5.95	.371	.287	.967
Participants							
1	3.93	4.11	4.17	9.93	.376	.282	.971
2	3.31	3.89	3.93	5.01	.408	.327	.93
3	3.79	4.26	4.04	10.88	.36	.291	.967
4	4.07	3.71	4.15	6.14	.143	.01	.655
5	3.35	3.62	3.89	15	.5	.447	.94
6	3.45	3.80	3.72	4.49	.41	.396	.928
7	3.63	4.15	3.92	4.07	.387	.289	.94
8	3.09	3.55	4.21	7.68	.4	.448	.89
9	4.10	3.90	4.15	7.87	.39	.335	.929
10	3.25	3.84	4.01	4.29	.4	.308	.942
11	3.28	3.87	4.19	9.66	.308	.261	.853
12	3.86	4.07	4.23	7.17	.32	.295	.965
13	3.47	3.89	4.16	5.19	.349	.266	.951
14	3.85	3.72	4.	4.14	.393	.249	.897
15	3.37	3.66	3.88	4.07	.332	.264	.865
16	3.85	4.22	4.17	13.30	.405	.32	.97

Note. Average parameters are based on data averaged over participants (not the average of individual parameters).

Table S2.3A

Parameter estimates from the composite list fit for emotional trials.

	Parameters			
	λ	β	δ	R^2
Average	3.88	4.23	.308	.97
Participants				
1	3.93	6.71	.364	.934
2	4.02	4.19	.381	.960
3	4.12	7.81	.362	.989
4	4.09	6.08	.092	.913
5	3.66	15.00	.512	.953
6	3.75	2.60	.373	.985
7	4.05	5.00	.402	.985
8	3.68	3.92	.404	.941
9	3.84	7.99	.377	.987
10	3.7	4.11	.478	.952
11	3.81	11.72	.356	.994
12	4.04	5.46	.288	.991
13	3.46	8.28	.362	.934
14	3.61	9.69	.393	.784
15	3.60	11.07	.421	.965
16	4.03	7.56	.355	.992

Note. Average parameters are based on data averaged over participants (not the average of individual parameters).

Table S2.3B

Parameter estimates from the composite list fit for neutral trials

	Parameters			
	λ	β	δ	R^2
Average	3.90	4.32	.296	.96
Participants				
1	4.05	11.97	.363	.994
2	3.71	4.98	.384	.948
3	4.10	6.68	.294	.987
4	4.00	4.82	.037	.743
5	3.63	15.00	.484	.975
6	3.66	4.52	.407	.949
7	3.89	4.51	.377	.960
8	3.63	6.72	.404	.913
9	4.01	10.14	.390	.931
10	3.68	4.65	.379	.952
11	3.78	10.12	.297	.924
12	4.05	7.25	.312	.989
13	3.83	5.26	.323	.968
14	3.93	2.97	.288	.924
15	3.64	4.06	.309	.885
16	4.07	13.99	.385	.997

Note. Average parameters are based on data averaged over participants (not the average of individual parameters).

Table S2.4A

Parameter estimates from the two-process model fit for emotional trials

	Parameters						
	λ_1	λ_2	β	δ_1	δ_2	γ	R^2
Average	0.15	0.062	4.6	0.297	0.508	-0.0119	0.984
Participants							
1	0.14	0.01	4.59	0.258	0.472	-0.0223	0.944
2	0.21	0.154	5.54	0.396	0.3	-0.00303	0.975
3	0.01	0.01	5.83	0.298	0.3	-0.017	0.989
4	0.043	0.082	7.79	0.1	0.3	-0.00121	0.992
5	0.33	0.094	6.06	0.416	0.753	-0.0003	0.939
6	0.35	0.01	3.38	0.416	0.518	-0.00007	0.976
7	0.44	0.031	6.36	0.429	0.534	-0.0108	0.974
8	0.22	0.029	4.58	0.409	0.915	0.0125	0.968
9	0.088	0.092	6.8	0.334	0.972	-0.0106	0.964
10	0.087	0.17	4.44	0.335	0.997	-0.038	0.969
11	0.01	0.051	7.2	0.31	0.668	-0.0138	0.974
12	0.19	0.01	7.86	0.283	0.501	-0.00295	0.984
13	0.22	0.01	9.79	0.346	0.543	-0.0372	0.946
14	0.025	0.26	4.24	0.296	1.57	-0.015	0.924
15	0.01	0.17	6.13	0.344	0.412	-0.0334	0.962
16	0.088	0.01	7.33	0.317	0.673	-0.01098	0.99

Note. Average parameters are based on data averaged over participants (not the average of individual parameters)

Table S2.4B

Parameter estimates from the two-process model fit for neutral trials

	Parameters						
	λ_1	λ_2	β	δ_1	δ_2	γ	R^2
Average	0.2	0.037	4.89	0.289	0.46	-0.006	0.985
Participants							
1	0.133	0.024	6.87	0.277	0.482	0.009	0.956
2	0.35	0.16	12.5	0.458	0.553	0.032	0.964
3	0.085	0.01	8.14	0.285	0.405	-0.0012	0.988
4	0.037	0.01	8.89	0.1	0.487	-0.0085	0.986
5	0.588	0.041	17.6	0.482	0.572	0.0094	0.977
6	0.085	0.01	9.72	0.485	0.698	0.0284	0.978
7	0.29	0.01	5.63	0.402	0.476	0.0004	0.979
8	0.28	0.12	8.43	0.405	0.819	-0.0436	0.947
9	0.35	0.01	6.88	0.32	0.407	0.0152	0.963
10	0.29	0.01	5.63	0.402	0.476	-0.013	0.979
11	0.1	0.01	10.0	0.286	0.393	-0.0165	0.994
12	0.19	0.01	9.53	0.304	0.405	0.003	0.994
13	0.32	0.01	7.68	0.32	0.444	-0.0211	0.974
14	0.22	0.01	3.68	0.304	0.464	-0.0035	0.959
15	0.2	0.058	4.51	0.303	0.525	-0.0185	0.957
16	0.01	0.01	11.6	0.375	0.3	0.0172	0.988

Note. Average parameters are based on data averaged over participants (not the average of individual parameters).

Supplement III

Stimuli Selection. We selected 100 images for each neutral category. Neutral categories were: a) furniture consisting of images such as sofas, desks, coffee tables, etc., b) kitchen utensils that included images of plates, pans, forks, etc. We selected the neutral images, mainly using a Google image search, such that the focal object in the image corresponded to an object listed in the database of category norms for words (Van Overschelde, Rawson, & Dunlosky, 2004).

The two emotion categories were: a) Disgust consisting of images that evoked disgust and b) fear including threatening fear evoking images. For each of those we initially selected around 115 images. Some of the stimuli used in this study were highly unpleasant and arousing images chosen from International Affective Picture System (IAPS) (Lang et al., 2005) which was a subset of the images used in our previous study (Mizrak, & Öztekin, 2016)[8]. To extend the stimulus set we again performed a Google image search and obtained other images with similar content. For the fear category, we used threatening stimuli such as snakes, guns, etc. For the disgust category, we used images such as rotten food, dirty toilets, insects, etc.

To select the final stimulus set for the emotion categories, we performed a pretest in which 18 participants who did not participate in the actual experiment rated all available emotional images. The pretest consisted of two sessions. In each session, participants were presented with images from both categories. There were 100 in the first and 130 images in the second session. We balanced the number of images that could fall into fear or disgust category for each session such that each session had equal number of images from the two categories. The experimental procedure for the two sessions was the same. The order of presentation of images was randomly chosen.

---

[8] Some of the IAPS image numbers that are used in the emotional stimuli set, fear category: 3005.1, 6190, 6212, 6370, disgust category: 3261, 9300, 9301, 2730.

In each trial, an image was presented for two seconds and participants were asked whether this image evoked fear or disgust for them. They were instructed to pick the emotion which they thought was more dominant by pressing ‘D’ or ‘F’ key on the keyboard. Next, they rated how disgusting they found this image from 1 to 9, and how fearful the image was from 1 to 9. Then, the image was presented one more time and participants were asked to give arousal and valence ratings to the image on a scale of 1 to 9.

When categorizing based on the predominant response, participants categorized 122 images as fearful and 108 images as disgusting. However, we discarded images for which the predominant response was selected with a ratio of 12:6 or less extreme for one of the categories. For instance, if an image had seven times fear response and eleven times disgust response, this image was not used in the experiment. For the final sample we selected 113 images for the fear category and 102 images for the disgust category. Images in the disgust category had significantly higher disgust ratings than fear ratings; t(203.93) = 16.27, $p < .0001$. Images in the fear category had higher fear ratings than disgust ratings; t(222.94) = 31.85, $p < .0001$. The images in the two categories did not differ in their arousal (t(208.96) = - 0.48, p = .63) and valence (t(208.96) = 1.53, p = .13). See Table S1 for mean and standard deviations of arousal and valence levels for both categories.

Table S1.1A
Means and standard deviations of arousal and valence levels for fear and disgust categories of emotional images

Category	Arousal (M)	Arousal (SD)	Valence (M)	Valence (SD)
Fear	5.35	2.5	2.75	1.91
Disgust	5.25	2.6	2.93	2.13

To further test whether participants were able to differentiate the two categories, we used Multi Voxel Pattern Analysis (MVPA) that can be utilized to classify cognitive states based on distributed patterns of neural activity in the brain (see Onal, Ozay, Mizrak, Oztekin, & Vural, 2016, for detailed results). We used the encoding data from 12 participants as the training set for the classifier and images that were presented at the retrieval phase as the test set. Our results showed that neutral and emotion categories were successfully classified as different categories by the classifier which had over 60% performance for each participant. The classifier was also successful in identifying sub categories of emotion and neutral images which were disgust, fear, furniture, and kitchen utensils (over 40% success for each participant).

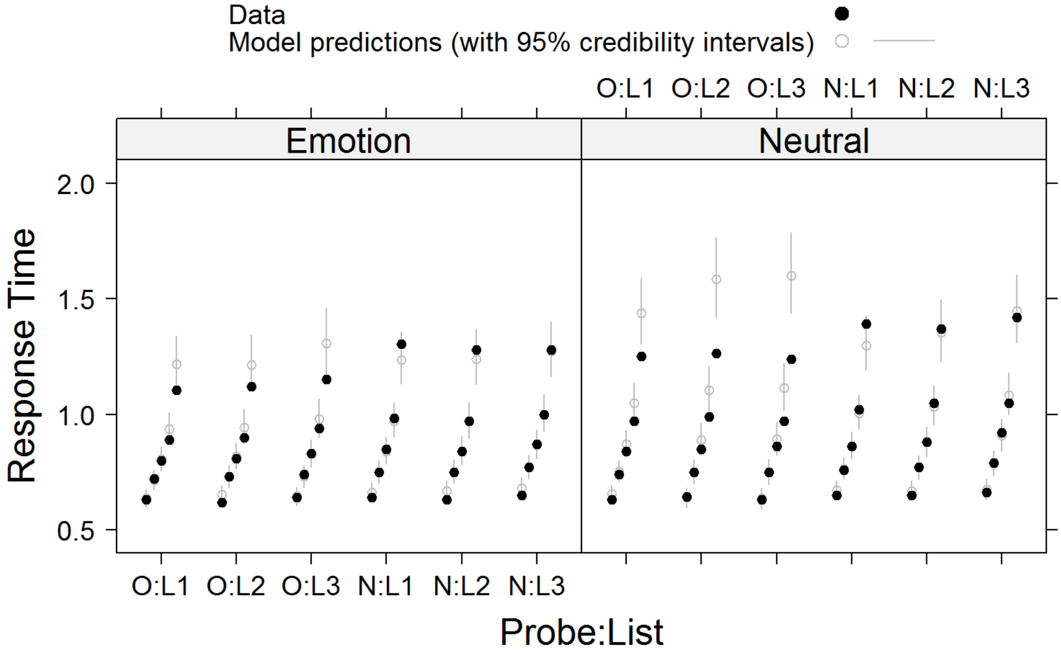

Figure S3.1. Quantiles of the observed response time distribution (in black) and corresponding model predictions (in gray, with 95 % credibility intervals) for correct responses. For each x-axis the five dots correspond to the 10 %, 30 %, 50 %, 70 %, and 90 % quantiles, respectively. With the exception of the slowest quantile (90 %) for old items, the model fits are good.

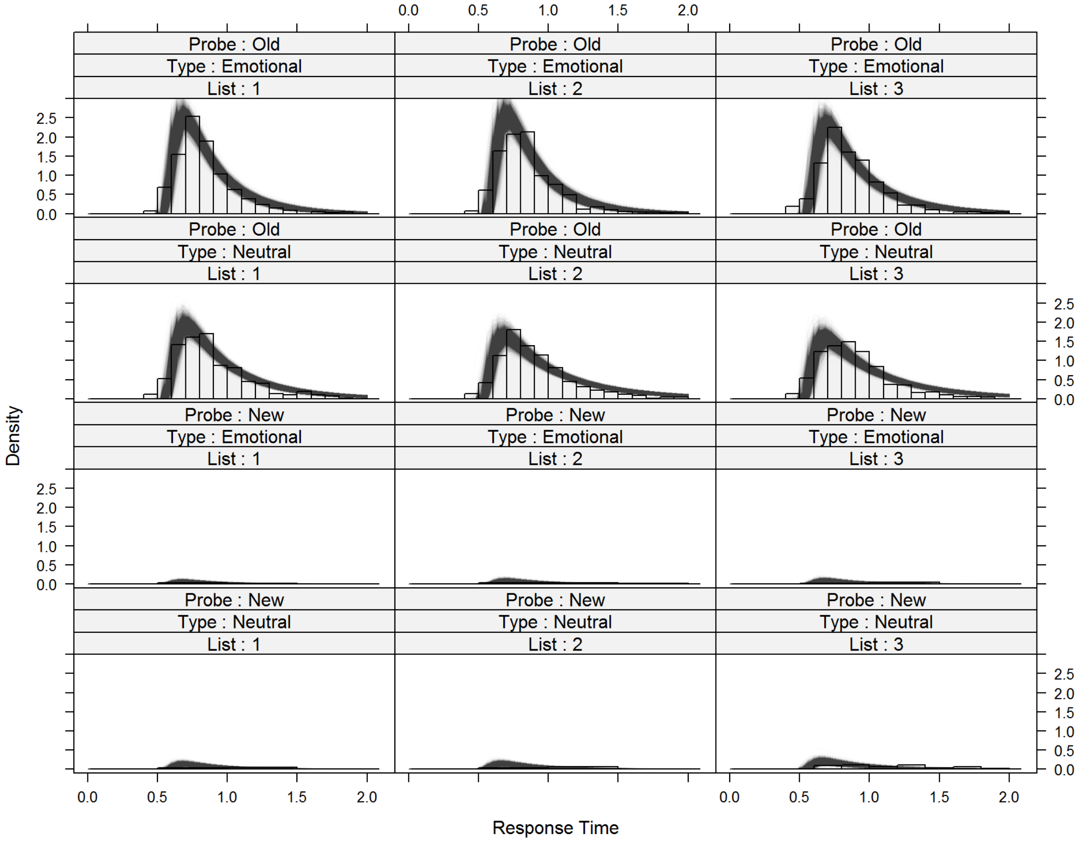

Figure S3.2. Full response time distribution of observed responses (histograms) and overlayed model predictions from the full posterior distribution for upper boundary (“old” response).

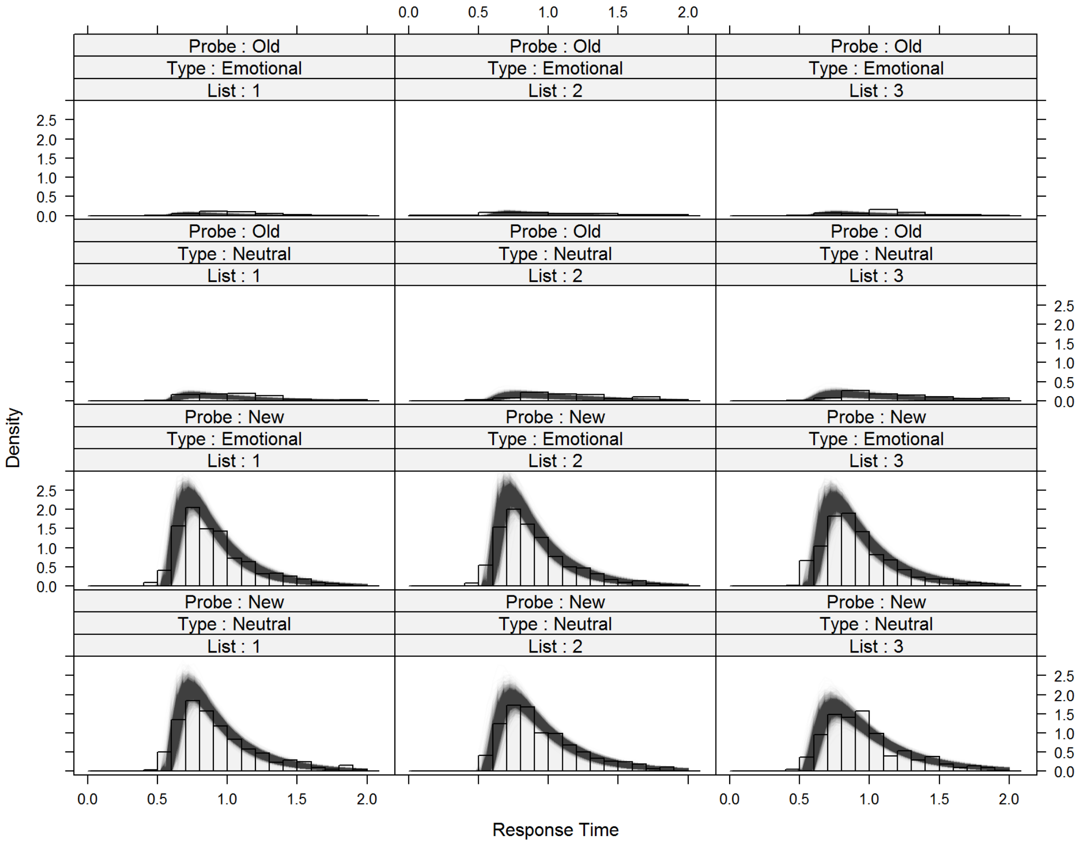

Figure S3.3. Full response time distribution of observed responses (histograms) and overlayed model predictions from the full posterior distribution for lower boundary ("new" response).