Prior Knowledge and Recognition Memory
a Computational Modeling Approach

Hellman, Johan

2013

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Prior Knowledge and Recognition Memory – a Computational Modeling Approach

Johan Hellman

LUND UNIVERSITY

To be publicly examined in Kulturens Auditorium, Lund University, on the 4:th of October 2013 at 10:00 for the degree of Doctor of Psychology.
Advisor: Sverker Sikström
Faculty opponent: Christopher Berry

Department of Psychology
ISBN: 978-91-7473-631-1 and 978-91-7473-632-8

Printed in Sweden by Media-Tryck, Lund University
Lund 2013

MILJÖMÄRKT
341903

CLIMATE
COMPENSATED
PAPER

REPA
A part of FTI (the Packaging and
Newspaper Collection Service)

Devoted to my family

Table of Content

Abstract

For more than a century, an immense interest has been devoted to the study of recognition memory, where a multitude of memory phenomena has been explained. Recognition memory is usually described with parsimonious measurement and statistical models, stemming from dual process theory and signal detection theory. In the present thesis, the most often used models of recognition memory are reviewed and compared to a novel implementation of the variance theory, abbreviated the VT (Sikström, 2001) in the account of frequency and familiarity effects, and a new model of item variability (the multidimensional signal detection theory, abbreviated the MSDT). The focus of the thesis lies on the effects of prior knowledge on recognition memory, investigated with behavioral, electrophysiological and modeling approaches.

In Study 1, a novel paradigm for measuring frequency and familiarity effects in recognition memory was introduced (the name paradigm), where recognition memory was tested on rare and common names that were famous and non-famous. The name paradigm was experimentally implemented in different conditions that provided a detailed description of fame and familiarity effects in recognition memory in four experiments. The study showed that pre-experimental knowledge both facilitates and impairs memory. Fame and frequency were selectively related to specific and non-specific semantic knowledge, where the former enabled retrieval of more and detailed information whereas frequency lacked such specificity at retrieval.

The second study elaborated on prior knowledge on recognition memory with the name paradigm by recording Event-Related Potentials, a method with which electrophysiological signatures of cognitive processes can be linked to experimental manipulations. More specific, it was investigated whether old/new effects previously related to familiarity and recollection, the FN400 old/new effect (Mecklinger, 2006) and the late positive component (the LPC, see Rugg & Yonelinas, 2003), respectively, would be selectively induced by frequency and fame, thereby linking the experimental variables to the two memory processes. Further, in a second experiment, the proposed link between familiarity and conceptual priming (Paller, Voss & Boehm, 2007) was investigated. The behavioral findings replicated those in Study 1, and the ERP analysis revealed that low frequent names elicited the FN400 effect, whereas fame to a higher extent than frequency gave rise to the LPC. Experiment 2 demonstrated that familiarity (i.e., the FN400) is insensitive to conceptual priming.

Study 3 provided a comprehensive account of fame and frequency effects by a novel implementation of the VT (Sikström, 2001). In two experiments the name paradigm was implemented in conditions where item, source and associative memory was assessed, which replicated the memory findings in Study 1 and 2. In the VT, fame was defined as a pre-experimental encoding of the stimulus. When a famous name was encoded the reinstatement of the item, based on previous experiences (prior to the experimental test) was associated with an increase in the specificity of the representation. This lead to an increase in net input to the underlying at retrieval, due to the high degree of similarity between the encoded and the retrieved item, and low degree of variability. Frequency, on the other hand affected the variability but not the magnitude of the net input, which resulted in lower memory performance.

In Study 4, a new model of item variability was introduced, the MSDT. The MSDT describes memory with three parameters, similar to the account provided by signal detection theory (SDT), but introduces non-linearity's to SDT, relies on binomial rather than normal latent distributions, and provides a multidimensional account of memory phenomena. The MSDT suggests novel predictions on changes in item variability as a function of attentional skill (i.e., ADHD versus healthy controls and varying degrees of attentional disabilities) as well as for the mediators in the differences in response variability in attentive and inattentive people. These predictions were tested on attentive and inattentive people, and provided augmented support for the model. The MSDT was conceptually and mathematically compared to the unequal-variance signal detection theory and the dual-process signal detection model (Yonelinas, 1994), and provided a more comprehensive account of the studied memory phenomena. Because attentive people yield a higher number of active nodes than attentive, and a lower variability in the activation threshold, the former group performs better and yields a higher ratio of new to old item variability than the latter. The MSDT also account for higher level of false alarms in inattentive than attentive, and suggests that the difference in new to old item variability is a result of increased new item variability relative that of old items. Further, the model provides a unified account of item- and response variability.

Organization Lund University Department of Psychology	Document name Doctoral Dissertation
	Date of issue 4 October 2013
Author(s) Johan Hellman	Sponsoring organization Swedish Research Council

Title and subtitle

Prior Knowledge and Recognition Memory – a Computational Modeling Approach

Abstract

For more than a century, an immense interest has been devoted to the study of recognition memory, where a multitude of memory phenomena has been explained. Recognition memory is usually described with parsimonious measurement and statistical models, stemming from dual process theory and signal detection theory. In the present thesis, the most often used models of recognition memory are reviewed and compared to a novel implementation of the variance theory, abbreviated the VT (Sikström, 2001) in the account of frequency and familiarity effects, and a new model of item variability (the multidimensional signal detection theory, abbreviated the MSDT). The focus of the thesis lies on the effects of prior knowledge on recognition memory, investigated with behavioral, electrophysiological and modeling approaches. In Study 1, a novel paradigm for measuring frequency and familiarity effects in recognition memory was introduced (the name paradigm), where recognition memory was tested on rare and common names that were famous and non-famous. The name paradigm was experimentally implemented in different conditions that provided a detailed description of fame and familiarity effects in recognition memory in four experiments. The study showed that pre-experimental knowledge both facilitates and impairs memory. Fame and frequency were selectively related to specific and non-specific semantic knowledge, where the former enabled retrieval of more and detailed information whereas frequency lacked such specificity at retrieval. The second study elaborated on prior knowledge on recognition memory with the name paradigm by recording Event-Related Potentials, a method with which electrophysiological signatures of cognitive processes can be linked to experimental manipulations. More specific, it was investigated whether old/new effects previously related to familiarity and recollection, the FN400 old/new effect (Mecklinger, 2006) and the late positive component (the LPC, see Rugg & Yonelinas, 2003), respectively, would be selectively induced by frequency and fame, thereby linking the experimental variables to the two memory processes. Further, in a second experiment, the proposed link between familiarity and conceptual priming (Paller, Voss & Boehm, 2007) was investigated. The behavioral findings replicated those in Study 1, and the ERP analysis revealed that low frequent names elicited the FN400 effect, whereas fame to a higher extent than frequency gave rise to the LPC. Experiment 2 demonstrated that familiarity (i.e., the FN400) is insensitive to conceptual priming. Study 3 provided a comprehensive account of fame and frequency effects by a novel implementation of the VT (Sikström, 2001). In two experiments the name paradigm was implemented in conditions where item, source and associative memory was assessed, which replicated the memory findings in Study 1 and 2. In the VT, fame was defined as a pre-experimental encoding of the stimulus. When a famous name was encoded the reinstatement of the item, based on previous experiences (prior to the experimental test) was associated with an increase in the specificity of the representation. This lead to an increase in net input to the underlying at retrieval, due to the high degree of similarity between the encoded and the retrieved item, and low degree of variability. Frequency, on the other hand affected the variability but not the magnitude of the net input, which resulted in lower memory performance. In Study 4, a new model of item variability was introduced, the MSDT. The MSDT describes memory with three parameters, similar to the account provided by signal detection theory (SDT), but introduces non-linearity's to SDT, relies on binomial rather than normal latent distributions, and provides a multidimensional account of memory phenomena. The MSDT suggests novel predictions on changes in item variability as a function of attentional skill (i.e., ADHD versus healthy controls and varying degrees of attentional disabilities) as well as for the mediators in the differences in response variability in attentive and inattentive people. These predictions were tested on attentive and inattentive people, and provided augmented support for the model. The MSDT was conceptually and mathematically compared to the unequal-variance signal detection theory and the dual-process signal detection model (Yonelinas, 1994), and provided a more comprehensive account of the studied memory phenomena. Because attentive people yield a higher number of active nodes than attentive, and a lower variability in the activation threshold, the former group performs better and yields a higher ratio of new to old item variability than the latter. The MSDT also account for higher level of false alarms in inattentive than attentive, and suggests that the difference in new to old item variability is a result of increased new item variability relative that of old items. Further, the model provides a unified account of item- and response variability.

Key words

Recognition memory, Episodic memory, Pre-experimental familiarity, Event-Related Potentials, Item variability, Receiver-Operating Characteristics, Computational Models

Language English	ISBN 978-91-7473-631-1 and 978-91-7473-632-8
Recipient's notes	Number of pages 214

Distribution by (name and address)

Signature Date

Svensk Sammanfattning

Tidigare Kunskap och Igenkänningsminne – ett Datamodellerings Perspektiv

Johan Hellman

Institutionen för psykologi, Lunds universitet, 2013

Studiet av igenkänningsminne har varit centralt i minneslitteraturen under mer än ett århundrade, vilket har resulterat i ökad förståelse för hur olika minnesprocesser fungerar och interagerar vid inkodning, lagring och framplockning av minnen. Igenkänningsminne är vanligtvis förklarat av mät-modeller och statistiska modeller vilka härstammar från två-process teori och signal detektions teori. I denna avhandling jämförs de mest populära modellerna av igenkänningsminne med en ny implementering av variansteorin (Sikström, 2001), för att förklara frekvens och familjaritets effekter, och en ny modell av konfidens- och responsvariabilitet, kallad MSDT (den multidimensionella signal detektions teorin). Huvudsyftet med avhandlingen är att studera effekterna av tidigare kunskap på igenkänningsminne, vilket sker med beteende-, elektrofysiologiska och modellerings metoder.

I studie 1 introduceras ett nytt paradigm för att studera frekvens och familjaritetseffekter där igenkänningsminne testades på vanliga och ovanliga namn vilka varierade i celebritet. Paradigmet testades med flera tekniker som ger en detaljerad beskrivning av celebritets och frekvenseffekter på igenkänningsminne i fyra experiment. Denna studie visade att igenkänning av kända och okända namn med hög och låg frekvens är kopplat till specifik och icke-specifik kunskap. Inkodning av namn på kända personer resulterade i högre minnesprestation jämfört med okända namn, samt framplockning av mer och detaljerad kontextuell information. Resultaten tolkades i enlighet med två-process teori, där de positiva effekterna av celebritet relaterades till minnesprocessen erinring (fri översättning av ”recollection”), och de negativa effekterna av frekvens relaterades till minnesprocessen bekantskap (fri översättning av ”familiarity”).

I studie 2 undersöktes effekter av tidigare kunskap på igenkänningsminne med metoden Event-Related Potentials (ERP), med vilken elektrofysiologiska signaturer av kognitiva

processer kan kopplas till experimentella manipulationer. Mer specifikt undersöktes huruvida de elektrofysiologiska signaturer som tidigare kopplats till processerna bekantskap och erinring, FN400 (Mecklinger, 2006) och den sena positiva komponenten (LPC, se Rugg & Yonelinas, 2003) respektive, induceras av frekvens och celebritet. I ett andra experiment undersöktes den påstådda kopplingen mellan konceptuell priming och bekantskap (Paller, Voss & Boehm, 2007). Beteenderesultaten replikerade fynden från studie 1, och ERP analysen visade att lågfrekventa namn gav upphov till FN400, emedan celebritet i högre utsträckning än frekvens inducerade den sena positiva komponenten (LPC). Experiment 2 visade att bekantskap (FN400 komponenten) inte var relaterad till konceptuell priming.

I studie 3 användes namnparadigmet för att replikera fynden från studie 1, och en ny implementering av variansteorin (Sikström, 2001) användes för att förklara dessa fynd. I två experiment studerades item minne, källminne och associativt minne med namnparadigmet, och fynden från studie 1 och 2 replikerades. I VT definierades celebritet som pre-experimentell inkodning av stimulus, där framplockning av tidigare inkodade stimulus ökar input till noder i nätverket eftersom det framplockade stimulus har hög koherens med tidigare inkodning av stimulus. Frekvens påverkar standardavvikelsen för input till nätverket, men inte storleken på input, eftersom hög frekvens är implementerat som ett högre antal pre-experimentella kontexter.

I Studie 4 introducerades en ny modell av konfidensdata (ROC kurvor) – MSDT. MSDT beskriver minne med tre parametrar, liknande signal detektions teori (SDT), men introducerar en icke-linjaritet i konventionell SDT, baserar sig på en binomialfördelning snarare än en normalfördelning av de underliggande distributionerna samt ger en multidimensionell förklaring av minnesfenomen. MSDT sätter upp nya prediktioner vilka testades på och bekräftades av data från personer med varierande nivå av uppmärksamhet (ADHD och friska kontroller). Modellen jämfördes med en variant av SDT i där de underliggande distributionerna antas ha olika varians, samt en två-process modell (Yonelinas, 1994). MSDT gav både en kvalitativt och kvantitativt bättre förklaring av empiriska data av ord igenkänning. Enligt modellen har uppmärksamma individer ett högre antal aktiva noder vid framplockning, och en lägre variabilitet i aktiveringströskeln jämfört med icke-uppmärksamma individer. Detta resulterade i högre minnesprestation och en lägre kvot av konfidensvariabilitet för studerade och icke studerade ord för uppmärksamma deltagare. MSDT ger även en förklaring till varför icke-uppmärksamma individer accepterar fler ostuderade stimulus som studerade jämfört med uppmärksamma individer, och föreslår att skillnader i variabilitet för studerade och ostuderade stimuli uppstår som ett resultat av variabilitet i distributionen för ostuderade stimulus, snarare än i den för studerade. Modeller ger även en enhetlig förklaring av konfidensdata och responsvariabilitet.

Acknowledgements

The work on this dissertation has been made possible thanks to the contribution of others, as well as the effort of both family and colleagues in the continued work.

I stand in great gratitude to my initial supervisor, Georg Stenberg. Thanks to you, I was involved in exciting and important research during my senior studies at Kristianstad University College. During the years we collaborated, you provided me both intellectual and amicable support, which made the continuing work with my dissertation, and my academic development possible. I would also like to thank Mikael Johansson, who introduced me to cognitive neuroscience and cognition. Your seminal work has been inspiring.

I want to thank my supervisor Sverker Sikström. You introduced me to the world of computational modeling, and, while I battled with the task to comprehend and learn the work within this realm, you provided me guidance. Your unconventional courses of action and seemingly boundless way of thought have been inspiring, and uncapped ideas and work that otherwise would not have been realized.

I also want to express my gratitude to Peter Juslin, for providing insightful and highly relevant critique and concrete suggestions for how to increase the quality of my dissertation.

Several other persons at the department of psychology have contributed to my academic work. I want to thank Per Johnson for you continued support in issues of both academic and personal nature. You have been a cornerstone in both harsh and flourishing times. Birgitta Abdon and Eva Henriksson, you both solved financial and administrative problems with ease. I also want to thank Lars Kindberg and Tommy Bauer for your arrangements and support with technical, administrative and practical issues. I thank Mats Dahl for help and support during my initial work at Kristianstad University Colleague and some periods at Lund University.

There are also individuals in my social and domestic circle who have provided support to the work of my dissertation.

I am grateful to my parents; Berton Augustsson and Magareta Fredriksson, for continuing support and help with baby-sitting and other issues. I am grateful for the continuing assistance and support from my parents in law, Elisabeth & Hans Hellman, for continuing assistance with my children and loving friendship. You have all provided me a great deal of help and support. Thanks to Magnus & Monica Hellman, and Lars Persson and Janina Hellman for interesting discussions, assistance with my children, and necessary leisure time. You have all been important for my work. I also want to thank Johan Mårtensson, Emelie Stiernströmer and Robin Hellerstedt, for interesting and inspiring discussions, practical advice and friendship.

However, of greatest importance for my work, and my life in general, is my wife: Åsa Maria Hellman, and our children, Love and Lo Hellman. The fulfillment of this dissertation had not been possible without you, neither had my personal or academic development been achieved if you had not been there. You have been my inspiration, my cornerstone, my primary reason for work and life. Thank you!

And, of cause, I must thank Ossie and Lissie Hellman, for you antihypertensive influence and ludicrous behavior that always results in laughs and love. And, to some extent, thanks to Scarlett Hellman, even though your domestic contribution has been more vexatious than calming, you bring joy to my life.

List of Papers

This dissertation is based on the following studies:

I. Stenberg, G., Hellman, J., & Johansson, M. (2008). The memorability of names and the divergent effects of prior experience. *European Journal of Cognitive Psychology, 20*(2), 312-345.

II. Stenberg, G., Hellman, J., Johansson, M., & Rosen, I. (2009). Familiarity or conceptual priming: event-related potentials in name recognition. *Journal of Cognitive Neuroscience, 21*(3), 447-460

III. Hellman, J., & Sikström, S. (Submitted for publication). The Variance Theory of recognition memory.

IV. Hellman, J., & Sikström, S. (Submitted for publication). The Multidimensional Signal Detection Theory.

Introduction

At a first glance, the study of recognition memory may seem mundane, because it simply encompasses the understanding of how we discriminate between experienced and not experienced events. However, we encounter persons and items in different contexts everyday, and despite processing of this massive amount of information, we can relatively accurately determine, for instance, when and where we last met a certain person. This makes the understanding of human recognition memory central for our knowledge of human memory. The topic has been thoroughly investigated, especially using *yes-no*, or *old-new* tasks where the testee is instructed to study a series of items (commonly words) and remember them for a subsequent memory test, where the old (studied) items are intermixed with new (unstudied) ones.

In early memory research, recognition was viewed as a simple form of recall but when it was shown that a single variable could affect recognition and recall processes differently, there was an increase in attention allotted to recognition. This interest has been maintained, in part due to the fact that recognition tasks are variable and pervious to the methodology used to study the neuropsychological basis of memory. Because recognition has been central to the study of memory, a variety of models have been proposed to account for several memory phenomena.

These models have been developed for two levels of understanding, dividing them into two specific classes of models, namely measurement and process models. On the one hand, measurement models describe recognition responses as a result of changes in response bias, due to differences in material, and task complexity. In such models, the processes delineating how memories are acquired, retained and retrieved are rarely described, whereas this dimension of memory is central in process models (Clark & Gronlund, 1996; Malmberg, 2008). Thus, a measurement model defines the limitations and concepts in the assessment of a certain behavior (what is possible?), and asserts, for example, that recognition is based on a continuous variable related to a specific measurement (i.e., memory strength). A process model tries to explain how this behavior is mediated, thus specifying how the measurement is generated from memory by computational modeling (Batchelder & Riefer, 1999).

A theoretical development that has made a large contribution to the understanding of memory is signal detection theory, hereon after abbreviated SDT (Egan, 1958), which has composed the foundation for subsequent memory models, such as global matching- (GM) and dual-process models. In short, SDT was developed to quantify the ability to distinguish between stimulus/signal and random patterns, or noise. Thus, the theory describes why a classifier is successful based on signal detection and a criterion that discerns signal from noise, and outlines how changes in the threshold affects performance. In contrast to signal detection models, dual-process theory, more specifically the dual-process signal detection model (abbreviated the DPSD), assume that recognition memory constitute two processes, familiarity and recollection, and assert that these two processes are necessary to adequately account for recognition memory (Yonelinas & Parks, 2007). Dual process models have been highly influential and has successfully accounted for a variety of experimental manipulations using several different methods and paradigms.

Purpose

The present thesis investigates two related issues of recognition memory. First, it aims to illuminate a limitation with current formal models of recognition memory, namely regarding the concept of pre-experimental familiarity (i.e., the knowledge about and/or familiarity with the test material used in an experiment), here operationalized as fame and frequency effects. In Study 1, a novel paradigm is introduced, which can be used to assess the contribution of fame and frequency on episodic recognition memory, implying how the two memory systems interact. In the second study, the interaction of episodic and semantic memory is studied with the name paradigm using event-related potentials (hereby after abbreviated ERP), which are neutrally generated voltage fluctuations recorded on the scalp that are time-locked to sensory, cognitive and motor activity. In the third study, the concept of prior knowledge is elaborated by implementing the variance theory, abbreviated the VT (Sikström, 2001) to account for the selective influence of the two variables on recognition memory. The VT is proposed as a formal account of fame and frequency effects in recognition memory, and is discussed in relation to SDT and the DPSD model.

Second, a limitation with SDT and the DPSD model is the lack of a formal, comprehensive account of item variability (i.e., the difference in old and new item variability in the latent distributions). Therefore, a new model is introduced in Study 4, the multidimensional signal detection theory (abbreviated the MSDT), which describes how and why variability in the new and old item distributions vary. Here, the MSDT accounts for differences in response variability, performance and item variability in attentive and inattentive people. The MSDT is discussed in relation to SDT and the DPSD model, and is both conceptually and mathematically compared to SDT and the DPSD model.

Theoretical Background

The theoretical division of human memory has been a subject of philosophical contemplation since Aristotle (Tulving, 1983), and given the work of Ebbinghaus (Ebbinghaus, 1885), and Tulving (Tulving, 1972, 1983, 2001), it has been a matter of scientific investigation rendering a large volume of publications both confirming and rejecting the concept of a divided memory system (Berry, Shanks, & Henson, 2008; McKoon & Ratcliff, 1986; Tulving, 2001; Yonelinas, 2001a). The early work on memory proposed a division of long and short term memory whereas Tulving suggested that human declarative memory constitute noetic and autonoetic experiences. The distinction between episodic and semantic memory has been supported by studies on patients with temporal lobe damage. For instance, these patients exhibit both anterograde and retrograde episodic memory impairments (Bayley, Hopkins, & Squire, 2006; Rosenbaum et al., 2008). Despite this severe memory impairment, medial temporal lobe (MTL) amnesiacs have preserved semantic knowledge from the premorbid period but lack the ability to retrieve episodic information. Further, patients that suffer from neocortical degeneration exhibit impairment in stimuli identification even though superordinate labels are provided (Hodges & Patterson, 1995). Albeit the MTL is affected, episodic memory is relatively spared (Graham, Becker, & Hodges, 1997). Thus, damages to the medial temporal lobe and areas inflicted in frontotemporal amnesia are selectively involved in episodic and semantic dementia. The distinction of episodic and semantic memory has resulted in thorough investigation of both types of memory, albeit along separate paths of inquire. As described in a recent review (Greenberg & Verfaellie, 2010), the interaction of the two systems has been devoted relatively limited attention, and the view on how interdependent they are varies. Despite this, the authors demonstrate that the two systems rely on each other at both encoding and retrieval.

Below, three aspects of recognition will be used to describe the interaction of episodic and semantic memory, regarding organization (i.e., interdependence of episodic and semantic memory), experimental manipulation (word frequency and fame) and theoretical and statistical assumptions (i.e., item variability).

Interaction of Episodic and Semantic Memory

It is well known that memory adapts in such a way that items likely to be encountered again will be better retained (Anderson & Schooler, 1991; Dennis, 1995), implying that episodic retrieval of items for which conceptual knowledge has been generated should be better than for semantically less known items. Several lines of investigations has been devoted to this notion, starting with, for example, the well known beneficial effect of depth of processing. That is, comparing different encoding manipulations reveal that a semantically elaborated study word is better retained than a perceptually encoded one (Craik & Lockhart, 1972). This manipulation (depth of processing) has influenced theories of memory that predicts that creation of new episodic memories benefits from established semantic knowledge (Tulving, 1983, 1985).

As mentioned above, Greenberg and Verfaellie (2010) described several studies implying an interdependence of the two memory systems, regarding interaction of episodic and semantic memory in both encoding and retrieval processes. For example, even though new semantic learning is impaired in amnesia (Verfaellie, Keane, & Johnson, 2000), certain experimental manipulations (stimulus variability and errorless learning techniques) can facilitate semantic learning (Stark, Gordon, & Stark, 2008; Tulving, Hayman, & Macdonald, 1991), and patients with early onset amnesia can retain semantic memories post morbidly (Kitchener, Hodges, & McCarthy, 1998; Vargha-Khadem et al., 1997). This indicates that episodic memory contributes to learning of semantic information, and other studies show that deficits in semantic memory are correlated with impaired episodic memory. For example, patients with dyslexia, semantic dementia and aphasia have exhibited impaired acquisition of new episodic memory (Graham et al., 2000; Kinsbourne et al., 1991; Ween, Verfaellie, & Alexander, 1996), indicating that impairment in episodic and semantic memory relates to each other, i.e., when semantic memory decreases, so does episodic memory performance.

The described research clearly indicates that episodic and semantic memory interacts during both encoding and retrieval, either because a selective compensation for learning in the absence of one of the two systems is necessary, or for a coherent decrease in performance for both systems. Findings that episodic memory performance co-varies in the presence of semantic memory impairment has been interpreted as inconsistent with the division of semantic and episodic memory. This view is supported by the notion that functional dissociations, or estimates of process purity, have been argued to be questionable when used to advocate different memory systems (see Toth & Hunt, 1999). Others argue that the distinction still has value for the understanding of selective memory impairments (Greenberg & Verfaellie, 2010).

Another example of the interaction of episodic and semantic memory is the hemispheric encoding/retrieval asymmetry model, or HERA (Habib et al., 2003; Tulving et al., 1994). According to HERA, encoding and retrieval asymmetrically involves different brain regions, more specifically, left prefrontal regions during encoding and right prefrontal regions during retrieval. Empirical support for the HERA model indicates that episodic and semantic memory interacts. Because the left prefrontal cortex is involved in processing of conceptual information (Badre & Wagner, 2002), it seems plausible that prior knowledge about a list item is activated during episodic encoding. On the other hand, activation of the right prefrontal cortex during retrieval has been related to retrieval mode (Lepage et al., 2000; Nyberg et al., 1995).

Experimental Manipulations

Frequency is a common variable in the memory literature because variations in frequency reflect differences in prior occurrence. Fame, on the other hand, is less studied even though famous people undoubtedly possess a special place in both semantic and episodic memory.

Word Frequency

The word frequency effect is a common finding in the memory literature, and infers that low frequent words are better remembered than high frequent ones (Chalmers & Humphreys, 1998; Reder et al., 2000). Commonly, levels of frequency affect the endorsement of new and old items differently. High frequency decreases hits but increases false alarms, whereas low frequency has the opposite effect, by increasing hits and decreasing false alarms. This regularity has been denoted the mirror effect, which has been a source of much debate and research (Glanzer et al., 1993; McClelland & Chappell, 1998; Shiffrin & Steyvers, 1997). It is important that formal models of memory can explain why there is a difference in memory strength for the easily remembered class of items when they are old and new (Hintzman, 1994). Or expressed differently, why do a condition leading to more hits (strengthening the memory trace), also affect responses to items not present on the study list? The word frequency effect has been explained in different ways, such as depending on differences in attentional resources (Glanzer & Adams, 1990b; Maddox & Estes, 1997), the number of retrieval processes (Joordens & Hockley, 2000), the number of contexts associated with the item (Dennis & Humphreys, 2001), and differences in encoding variability (McClelland & Chappell, 1998). Further, others argue that differences in commonness of orthographic features and normative versus letter frequency in words (Malmberg et al.,, 2002; Malmberg & Murnane, 2002; Shiffrin & Steyvers, 1997) or differences in variability caused by high and low frequent items (and features) in the connectionist memory system (Sikström, 2001; Sikström, 2004) mediates the mirror effect. Some models account of the mirror effect is delineated under the heading Memory models – a selective review, namely that of SDT and the DPSD.

Fame Recognition

Frequency is a highly relevant variable of prior knowledge because it represents knowledge of a statistical regularity of the world, guiding us in assessment of novelty and familiarity. Fame (i.e., a name or face of a famous person), on the other hand, has rather different implications for memory, because exposure to a famous name or face evokes conscious processing of a specific representation of the item that has been encoded at a previous occasion.

A famous name elicits retrieval of personal autobiographical experiences specifically related to that person, yielding facilitating effects in different memory tasks. For example, Westmacott and Moscovitch (2003) investigated whether variations in autobiographical significance for famous names resulted in different levels of episodic memories (i.e., more or less episodic information) in adult and old participants. Indeed, famous persons that were associated with retrieved events in the participants life was better remembered in an episodic memory test, and lead to improved semantic memory. Another example can be found in a recent study on frequency band analysis of data on recognition of famous and non-famous faces (Zion-Golumbic, Kutas, & Bentin, 2010). The authors observed increases in theta and alpha responses (i.e., reflecting episodic memory) during study of famous faces but not at retrieval, and gamma activity decreased from study to test, but was larger for famous as compared to non-famous faces. The result was interpreted as that gamma activity reflects activation of specific representations related to a certain famous person, which is beneficial for the creation of a novel episodic memory. A non-famous face, on the other hand, rather evokes subordinate representations, such as gender, unrelated to a specific person or characteristic.

The effect of fame on cognitive processing may be related to a constitutional principle of memory, namely the encoding specificity principle, or the ESP (Tulving & Thompson, 1973). According to the ESP, memory performance is highly influenced by consistency of contextual information of the event during encoding and retrieval. Thus, encoding operations determine what is stored in memory, and defines which retrieval cues are effective for subsequent access to what is stored. It can be argued that retrieval of a famous name evokes representations similar to those evoked at a previous encounter to the item, because a famous name is strongly associated with specific autobiographical information (Westmacott & Moscovitch, 2003), and this associated contextual information is strengthened by accumulated retrieval of the name in the past. On the other hand, prior knowledge may distort the beneficial effect described by the ESP for some items (i.e., non-famous, especially high frequent items), because the associated contextual information (i.e., the study list) is difficult to specifically relate to the item.

The described research show that semantic processing of fame is beneficial for episodic recognition, but only describes this effect in reference to autobiographical relevance. As will be showed in Study 1, 2 and 3, the facilitating effect of fame (relative frequency) may relate to a dividing line within semantic memory irrespective of autobiographical associations, making it relevant for the understanding of the organization of memory and significance for episodic recognition.

Item Variability

A central question in the present thesis is how to accurately account for item variability. Item variability refers to differences in recognition judgments to old and new items in a recognition test, that is, why the variability for studied and unstudied items differ, and how this difference can be accounted for. A popular technique for the investigation of episodic memory, and comparison of episodic memory models is the receiver-operating characteristic technique, or ROC curves (Macmillan & Creelman, 2005). An ROC curve can provide a measure of item variability, reflected by the slope of the z-ROC. The z-ROC measure is described in detail under the heading Measurements.

The ROC technique has been used extensively to compare models of recognition memory (Yonelinas & Parks, 2007) because it demands a model to describe the relation between accuracy and response bias and sets constraints on the model as it must account for several points along the ROC curve. Because an ROC reflects both isosensitivity and isobias, an ROC model can derive estimates of recognition performance undistorted by the response criterion used at a certain recognition judgment, and can therefore differentiate the influence of accuracy and response bias of an experimental manipulation. Investigation of recognition memory with ROC curves have influenced and guided work that has resulted in the current most influential model of recognition memory, namely SDT. Previous to, and parallel with the establishment of SDT, models described recognition as a threshold process (i.e., threshold theories), which predicted linear ROC curves. However, because the majority of all published ROC curves during the last four decades (Kinchla, 1994; Murdock, 1974) have been curvilinear, the threshold models was rejected in favor of SDT.

The study of item variability is important because it delineates several aspects of recognition (i.e., correct and incorrect memory judgments over a range of bias for a high number of materials and experimental manipulations) and provides an informative index of memory used to test computational models of recognition memory. Even though the z-ROC measure has been used extensively in the memory literature, the basic question why the old and new item variability differs is still unresolved (see Koen & Yonelinas, 2010; 2013).

Electrophysiological basis of recognition

Usage of the ERP method has resulted in a large contribution to the understanding of recognition memory, of which only a subpart will be mentioned here, regarding episodic memory retrieval (the method is further described under the heading Measurements). Within the realm of recognition memory, it has been demonstrated that correctly remembered old items elicit more positive-going ERPs than correctly rejected new items, a phenomenon that has been called the old/new effect (Neville et al., 1986; Rugg & Doyle, 1992; Wilding & Rugg, 1996). Because the effects do not occur for un-responded old items, it seemingly reflects successful retrieval from a previous event rather than stimulus repetition, or the mere "old" response. The old/new effect has been subdivided into several subcomponents as a function of different spatiotemporal distributions (Friedman & Johnson, 2000; Mecklinger, 2000). Bellow, two old/new effects related to different qualities of episodic retrieval (familiarity and recollection), or memory processes, is outlined.

Signatures of Familiarity and Recollection

The proposed memory processes familiarity and recollection has been differentially related to electrophysiological signatures, or old/new effects. Familiarity has been associated with an enhanced positivity for old relative new items during the time window of 400-600 ms. post stimulus onset, called the FN400. The effect is commonly observed over mid-frontal scalp distributions that extends to left and right frontal areas as well as central midline regions (Friedman & Johnson, 2000; Rugg & Curran, 2007; Mecklinger & Meinshausen, 1998). Examples indicating the link between familiarity and the FN400 are that the effect is insensitive to levels-of-processing manipulations (Rugg et al., 1998) and is similar for words and pseudowords (Curran, 1999).

Recollection, on the other hand, has been associated with a later component, known as the late positive component (the LPC), that occurs approximately at 400-800 ms. post stimulus with widespread scalp distribution (Curran, 2000; Rugg et al., 1998; Van Petten, Senkfor, & Newberg, 2000; Wilding, 2000; Wilding, Doyle, & Rugg, 1995). The effect is associated with recollection because it is larger for correct than incorrect source memory judgments (Wilding, 2000; Wilding & Rugg, 1996) and for deep as opposed to shallow studied items (Paller & Kutas, 1992). Another interesting difference between the two effects, that has been reported (Woodruff, Hayama & Rugg, 2006; Yu & Rugg, 2010), is that the FN400 increases gradually with recognition confidence, whereas the LPC only emerges for high confident responses. This has been interpreted as consistent with a dual-process perspective, more precisely, the DPSD model.

Duarte (et al., 2004) recorded ERPs during both encoding and retrieval while participants completed two different tasks each consisting of two consecutive study blocks, followed by a test block encompassing all studied items. Items were either judged in terms of animacy or manipulability, and in the subsequent test block, each item was judged as remembered, known or new (i.e., the Remember/Know test, described under the heading Measurements), and identified as coming from one of the two tasks (i.e., the source memory task, see Measurements). Subsequent familiarity and recollection based recognition (i.e., the FN400 and the LPC based on differences in remember/know and source memory responses) was preceded by different scalp topography and time course at encoding. That is, not only was the results consistent with the hypothesis that familiarity and recollection is related to different ERP components, the finding also suggest that familiarity and recollection are dissociated at both encoding and retrieval. In a more recent study, Addante (et al., 2012) recorded ERP on amnestic patients with deficits in recollection but not familiarity (an impairment documented with behavioral measures in previous studies, see Aly et al., 2010; Diana, Yonelinas & Ranganath, 2008), and was compared to healthy controls. The amnestic patients successfully discriminated old from new items but performed near chance in a source recognition task. Controls exhibited both the FN400 and the LPC, whereas the latter effect was absent in the amnestic patients.

Paller and colleagues advocates a different view (Yovel & Paller, 2004; Voss & Paller, 2009; Paller, Voss & Boehm, 2007; Voss, Lucas & Paller, 2010; Voss, Lucas & Paller, 2012; Paller, Lucas & Voss, 2012), namely that the dissociation of familiarity and recollection cannot be based on the association between each process and a distinct old/new effects. On the one hand, Yovel and Paller (2007) did not observe the FN400 for familiarity based recognition, and both types of memory experience were associated with bilateral, parietal brain potentials, where the effects of familiarity was lower than that of recollection. Recently, it was proposed that the FN400 only is correlated with familiarity in restricted circumstances, where implicit memory co varies with familiarity (Paller, Lucas & Voss, 2012). It should be noted that this view differs from that in Yovel and Paller (2007), or at the time for Study 2 in the present thesis, where the relation between familiarity and conceptual priming is investigated.

Memory Models – A Selective Review

To delineate the current knowledge on how pre-experimental familiarity interacts with memory, two different theoretical frameworks will be considered, which currently divide the memory literature, namely single and dual process theories of recognition memory. In the current thesis, SDT represents the former and the DPSD model the latter (Yonelinas, 1994).

Signal Detection Theory (SDT)

SDT stresses the uncertainty in a recognition situation, and provides an account of recognition memory using a precise language as well as graphical notations for the analysis of memory decisions. According to standard SDT, memory is described by two equal-variance Gaussian distributions (abbreviated the EVSD model), reflecting two random variables of signal and noise along a familiarity continuum. Signal reflects items that have been studied during the encoding phase of the test, whereas noise represents new items, which are intermixed with the old items in the test phase of the experiment. Participants separate old and new items by an assessment of item familiarity compared to a subjective response criterion.

When the familiarity value of an item exceeds the response criterion, it is either correctly endorsed (a hit) or incorrectly accepted as old (a false alarm). An item with familiarity not reaching the response criterion is either correctly rejected as new (a correct rejection) or incorrectly dismissed (a miss). Based on empirical data, the EVSD has been rejected and replaced with the unequal variance SDT (abbreviated the UVSD model), where the old item distribution is assumed to have a higher variability than that of the new items, because the majority of recognition memory studies during the last centuries have reported a flatter old than new item distribution (Mickes, Wixted, & Wais, 2007; Ratcliff, Sheu, & Gronlund, 1992b; Yonelinas & Parks, 2007).

Because the proportion of responses for the old and the new item distributions equals 1, hits and false alarms provides sufficient information about the participants recognition responses. The old item distribution has a lower variability than that of the new items, and the latter is normalized so that it equals 1.

Therefore, recognition memory can be described by the parameters d' (performance), C (the response criterion) and the old item variability. If the response criterion is moved, the proportion of hits and false alarms are changed. For a higher, conservative response criterion, the proportion of hits and false alarms will decrease, whereas a lower, liberal response criterion will result in a relatively high endorsement of both new and old items. The criterion, then, constitute an interpretation of response bias, and performance (d') equals the distance between the z-transformed hit and the z-transformed false alarm, which corresponds to the distance between the mean of the new and the old item distribution.

Signal detection theory has been influential for the understanding of memory for two reasons. First, the framework has suggested how to define and operationalize sensitivity and bias, and study these aspects of memory independently. Second, signal detection theory proposes that recognition memory can be described as a single random variable, often denoted memory strength, or familiarity (Macmillan & Creelman, 2005; Snodgrass & Corwin, 1988).

Familiarity Effects

Signal detection theory states that common words are more familiar than rare words, but that the unfamiliar character of a low frequent word increases the strength for this item in an accumulative manner at encoding. Thus, the latent distribution for new high frequent words is positioned more to the right than that of a new low frequent word, whereas the opposite pattern occurs for old items. That is, the distribution for an old rare word is to the right of the distribution of an old common word. This is why the mirror effect occurs; hits for low frequent word exceed that for high frequent words, whereas false alarms is higher for common than rare words. However, this only describes the statistical regularity of the frequency effect, and provides no psychological account of the phenomenon. Several proposals have been made regarding what psychological processes contribute to these frequency effects in the memory literature. Two common suggestions are levels of processing (Hintzman, 1988), implying that rare words are assumed to be more elaborated at encoding than common words, and that low frequent words receive more attention than high frequent ones (Glanzer & Adams, 1990a; Malmberg & Nelson, 2003).

ROC Interpretation

According to the UVSD, the increased old item variability results in an asymmetrical ROC in relation to the diagonal in p-space, with a corresponding z-slope below 1.0. Thus, performance is described by two parameters, d' and z-slope. The dissociation between performance and asymmetry has been investigated in several studies (Glanzer et al.,, 1999; Ratcliff, McKoon, & Tindall, 1994), that is, how recognition accuracy co varies with symmetry. This has been studied with several manipulations, such as word frequency (Arndt & Reder, 2002; Glanzer & Adams, 1990a; Glanzer et al., 1999), word concreteness (Glanzer & Adams, 1990a), list length (Gronlund & Elam, 1994; Ratcliff, McKoon, & Tindall, 1994), divided attention (Yonelinas, 2001b) and aging (Glanzer, Hilford, & Kim, 2004; Healy, Light, & Chung, 2005; Howard et al., 2006). According to some, accuracy and symmetry are inversely related, because high performance decreases the z-slope as higher variability in the old as compared to the new item distribution leads to higher performance. Others argue that an increase in accuracy is related to a relatively low decrease in z-slope (Hirshman & Hostetter, 2000). A different view has been proposed by Ratcliff and colleagues (Ratcliff, McKoon & Tindall, 1994; Ratcliff, Sheu, & Gronlund, 1992a), namely that symmetry is relatively constant irrespective of experimental manipulation and accuracy, this constant value being approximately 0.80.

The Dual Process Signal Detection Model (the DPSD)

During the last three decades, there has been a periodically fervent debate on the nature of recognition memory (Curran, 2000, 2004; Dunn, 2004, 2008; Heathcote, 2003; Smith & Duncan, 2004; Squire, Wixted, & Clark, 2007), regarding whether recognition memory constitute one or two variables (i.e., memory strength or familiarity and recollection). The dual-process perspective has gained support based on behavioral (Yonelinas & Parks, 2007), electrophysiological (Rugg & Curran, 2007) and imaging studies (Aggleton & Brown, 2006; Diana, Yonelinas, & Ranganath, 2007; Qin et al., 2009; Rugg & Yonelinas, 2003; Skinner & Fernandes, 2007). Accordingly, familiarity is based on non-specific, varying level of memory strength, consciously accessible as a feeling of familiarity for the presented item. Recollection, on the other hand, makes possible retrieval of contextual information, leading to a reconstruction of the episode in which the presented stimulus was previously experienced. Proponents of the single process perspective have emphasized the concept of parsimony, and when direct support for the two processes can be questioned, a more constrained model is to be preferred. Further, several studies have demonstrated weaknesses with interpretations of results from some of the paradigms commonly used to assess the contribution of familiarity and recollection, such as the remember-know paradigm (i.e., that the remember-know distinction really reflects differences in memory strength, see Dunn, 2008; Wixted, 2009; Wixted & Mickes, 2010) and the dual-process interpretation of ROC data (Ratcliff, McKoon, & Tindall, 1994; Wixted, 2007). Further, in the most influential dual-process model (Yonelinas, 1994), recollection is conceptualized as a threshold process, which has been questioned (Mickes, Johnson, & Wixted, 2010; Mickes, Wais, & Wixted, 2009; Slotnick, 2010; Slotnick & Dodson, 2005), opening for the possibility that only one process operates at recognition. Another criticism of dual process theory is that the division of recognition memory reflects different levels of memory strength, which should be regarded as a continuum from weak to strong memories rather than a division into two different memory processes (Squire, Wixted, & Clark, 2007; Wais, 2008).

Several dual-process models have been introduced (Mandler, 1980; Reder et al., 2000; Rotello, Macmillan, & Reeder, 2004; Yonelinas, 1994) of which the most influential is the DPSD model (Yonelinas, 1994).

According to the DPSD model, familiarity is as a signal detection process, indexed by d′, meaning that recognition by familiarity encompasses an assessment of memory strength, where accuracy is determined by changes in the response criterion. Thus, familiarity is described as a function of d' and C. Therefore, familiarity results in both correct and incorrect responses, as well as guesses. Recollection is conceptualized as a threshold process. An item is endorsed as old if it exceeds the recollective threshold, or if the familiarity assessment exceeds the response criterion in the absence of recollection. New items will only be endorsed if they are familiar, but will not reach the recollective threshold. Therefore, recollection only contributes to correct recognition, and does not generate false alarms.

Familiarity Effects

In comparison with single process models, the DPSD takes on a different perspective to explain frequency effects (Yonelinas, 2002). In general, low frequency words tend to induce recollection responses to a higher degree than high frequency words, whereas high frequency often result in familiarity assessments without a recollection of the study episode. Familiarity, on the other hand, is continuous, and results in both hits and false alarms. This is why a low frequent item exhibits higher levels of hits than false alarms (due to disproportionate use of recollection), and high frequent items are associated with more errors and fewer hits (a combination of familiarity and recollection).

By describing recognition memory as constituting both a graded process and a threshold process, Yonelinas introduced some novel aspects of the understanding of ROC data, which gained the model support.

ROC Interpretation

According to the DPSD model, familiarity and recollection have different effects on item variability, and the ROC account is therefore somewhat different from that of the UVSD. Because recollection is described as a threshold process, no incorrect responses are made when recollection governs recognition, resulting in a high proportion of high confident hits (and no false alarms). Familiarity, on the other hand, leads to both high and low confident responses and guesses. Therefore, the combination of the two retrieval processes results in an increase in the old item variability, and consequently, a lowered z-slope (Yonelinas, 2001b; Yonelinas & Parks, 2007). Because recollection increases ROC asymmetry and familiarity does not, z-slope and performance can be dissociated. Thus, conditions in which recollection drives recognition to a higher degree than familiarity leads to a flatter ROC in p-space, and a lower z-slope.

Limitations with the Current Theoretical Frameworks

Even though SDT provides an intuitively plausible and viable approach to understand recognition memory, there are some crucial limitations with the UVSD account. It states that recognition memory should be described by two distributions with unequal variability, due to empirical findings in support of the unequal model (Mickes et al., 2007; Ratcliff, Sheu, & Gronlund, 1992a; Yonelinas & Parks, 2007). However, it does not explain why this unequal variability occurs. Instead, proponents of SDT often rely on the verbally formulated encoding variability hypothesis (the EVH), as described by Wixted (2007). The EVA only asserts that the unequal variability is determined by item familiarity, and provides no formal account of the unequal variability. Further, the EVA provides no elucidation on how encoding variability affects new items, which is problematic because prior knowledge about the test items is known to affect recognition memory, as described in the introduction. The EVH has also been empirically tested and questioned. Koen and Yonelinas (2010) tested and rejected the EVH as accounting for unequal variability of the latent distributions. In commentary articles (Jang, Mickes, & Wixted, 2012; Starns, Rotello, & Ratcliff, 2012), the study was criticized on both statistical and methodological grounds. Rather than taking a clear stance on the outcome of this debate, where plausible arguments were provided by both the proponents of the EVH and the DPSD (Koen & Yonelinas, 2013), it may be argued that the EVH is generic and rather unclear.

The DPSD provides a formal account of item variability, where the contribution of familiarity and recollection has opposing effects on recognition memory, and thereby offer a theoretical account for changes in the z-slope. However, the model assumes that recollection is a threshold process, which is indispensable for the account of memory, and that familiarity and recollection are qualitatively independent processes. These assumptions have been questioned (Dunn, 2004; Mickes et al., 2009; Ratcliff, Van Zandt, & McKoon, 1995; Slotnick, 2010; Wais, Mickes, & Wixted, 2008).

A final limitation with the described ROC accounts is the fact that they merely encompass variables that generically influences memory performance, and commonly is used solely to compare memory models. However, an important goal with the present thesis is to show that it is possible to make a more elaborate interpretation of ROC data.

Measurements

In the present thesis, different paradigms, techniques and procedures are used to study episodic recognition, the contribution of prior knowledge on recognition memory, and item variability. These are shortly described here.

Remember-Know

The remember-know paradigm was introduced by Tulving (1985), with the purpose to measure the selective contribution of semantic and episodic memory in recognition tests. When the participant carries out a recognition test, each item response is accompanied with a remember/know response, where the testee declares whether the item is remembered (i.e., based on a re-experience of the study episode) or known (i.e., based on a familiarity assessment of the item with no re-experience of the study episode). Thus, the two response alternatives were thought to reflect episodic and semantic memory. According to dual-process theory, recollection of the study item and its context results in a remember response, whereas a know response reflects familiarity (Yonelinas, 2002). The use of the paradigm as suggested by dual-process theorists has been criticized, because several studies indicate that remember and know responses likely mirrors different levels of confidence, rather than different retrieval processes (Dunn, 2004, 2008; Wixted & Stretch, 2004). The remember-know paradigm is used in Study 1 (experiment 2).

Source Memory

The source memory framework provides a means to estimate the ability to assort relevant retrieved contextual information and the test item, a process made possible by source monitoring (Johnson, Hashtroudi, & Lindsay, 1993).

When a recent or remote event is remembered, different facets of that experience (i.e., the color of an encoded word or spatial location of a studied item.) are connected during encoding. Studying source memory involves investigating the effect of the binding of these different facets on memory (Johnson, 2006). Thus, remembering a source of an item demands successful retrieval of differentiated contextual information relevant for the item and accurate monitoring of these representations. The source memory framework constitutes an adequate assessment of how manipulation of pre-experimental familiarity affects episodic memory because the source memory ability and general recognition performance can be dissociated. According to dual-process theory, source memory performance reflects the contribution of recollection (i.e., successful retrieval of contextual information from the study event), whereas item memory is based on both familiarity and recollection. Thus, source memory tests can be used to study the selective contribution of the assumed memory processes. The source memory framework is used in Study 2 (experiment 1 and 2).

Associative Memory

The paired associates test measures associative memory, which reflects retrieval of a test item induced by retrieval of an associated item or context. Specifically, paired associate tests estimates the ability to remember two items that are presented in combination at study. In Study 2, (experiment 2) participants encode a pair of names at encoding, and are tested on old (a subset of the pairings presented at encoding), recombined (e.g., one old and one new item) and new combinations at test. By comparing performance on these different pairings, the contribution of the experimental manipulation on episodic memory can be assessed in detail. Because correct responses to recombined items demand a re-experience of the study episode, which a correct item response does not, dual-process theory suggest that paired associate test can be used to estimate the differential contribution of familiarity (correct/incorrect old and new pairings) and recollection (correct old/new and recombined pairings) (see Yonelinas, 2002).

In the studies included in this thesis, both source- and associative memory tests are used to investigate how different facets of episodic memory and pre-experimental familiarity affects recognition, and how these effects can be accounted for with the VT.

Receiver-Operating Characteristic (ROC) curves

ROC curves are computed by plotting hits and false alarms over different levels of bias, recorded by asking participants to rate their old-new responses with confidence, ranging from 0-100% with at least 3 ranges for both old and new items. Hits and false alarms are then accumulated over each confidence range, and are plotted as ROC curves. To estimate the variability of the underlying familiarity distributions of recognition memory, the z-slope of the ROC is often used. By plotting z-transformed hits against z-transformed false alarms across over the confidence intervals, the z-slope is computed (Egan, 1958; Glanzer et al., 1999; Yonelinas & Parks, 2007). A z-slope at 1 reflects equal variability of the underlying familiarity distributions, whereas an increase in variability of the old items pushes the z-slope below 1, or in rare cases, above 1.

ROC curves are computed in Study 1 (experiment 3 and 4), in Study 2 (experiment 1) and in Study 4.

Event-Related Potentials (ERP)

ERPs are electrophysiological responses embedded in a background electroencephalogram (EEG), but the EEG is assumed unrelated to the event of interest, which is why the ERPs can be extracted from the EEG by averaging the signal over a sufficient number of trials. When a certain number of trials are averaged over repetitions where the same type of event occurs, background noise is reduced in relation to the ERP, and the ERP, characterized by positive and negative peaks (so called components) related to the event of interest, can be identified. The ERP is related to sensory, cognitive and motor processes in terms of polarity, latency and scalp distributions, and are correlated with experimental variables and behavioral responses (Luck, 2005).

The ERP method has a low spatial but high temporal resolution, making it possible to reveal aspects of time course of cognitive processing, such as to determine necessary time periods for processing different types of items (Rugg & Coles, 1995). Further, the method allows complete randomization of experimental conditions across trials and averaging across combinations of item type (old and new items) and different behavioral estimates (performance, response-type etc.). The ERP technique is used in Study 2, experiment 1.

Empirical Studies

The overarching goal of the present thesis was two folded; first, to increase the understanding of how episodic and semantic memory interacts, and to show how this affects recognition memory (i.e., performance, response bias, item variability and electrophysiological correlates of recognition memory), and further, to provide a formal account of these effects. Another goal with the thesis is to provide a novel account of item variability, which overcomes the rather serious limitations in the current frameworks for item variability (SDT and the DPSD model). The four included studies are briefly described below, accompanied with a short delineation of the VT (Study 2) and the MSDT (Study 4).

Study 1 – Conceptual Influences on Episodic Memory

The limitations with the theoretical framework of pre-experimental familiarity motivate three studies in the present thesis (Stenberg, Hellman & Johansson, 2008; Stenberg et al., 2009; Hellman & Sikström, submitted). First, as was described in the review of the memory models, there is no consistent understanding of both frequency and fame effects, which was also pointed out by Malmberg (et al., 2002). The aim of Study 1 was to investigate the effects of fame and frequency on recognition memory performance, and shed light on the contradiction generated by theoretical predictions. More specifically, on the one hand, high frequency should increase memory performance because a common name would be better retained because they are most likely to encounter again, as compared to low frequent names (Anderson & Schooler, 1991). On the other hand, rare words are usually better remembered than common ones, as described by several models. Thus, the word frequency effect introduces a problem for the generalization that pre-experimental knowledge enhances memory performance. Further, because previous studies on fame effects mainly focus on autobiographical significance, where fame is viewed as directly related to episodic encoding, the present thesis investigates fame in a broader perspective in purpose to provide a more detailed account of fame effects.

To study the interaction of episodic and semantic memory, a new paradigm was developed - the name-paradigm (Stenberg, Hellman & Johansson, 2008). The participant is presented with and instructed to remember famous names, both frequent (e.g., using English equivalents; Tom Jones, Gordon Brown) and infrequent (e.g. Gwyneth Paltrow, Javier Bardem), as well as non-famous names, also frequent (e.g. John Smith, Jane Cooper) and infrequent (e.g. Sebastian Weisdorf, Brogan Kincaid).

The name paradigm was implemented in 4 experiments, where different techniques were used to elaborate on the beneficial and detrimental effects of fame and frequency on recognition memory. The first three experiments enrolled 47, 35 and 28 participants, and in the fourth experiment, 24 participants were recruited (conducted in an ERP recording context, electrophysiological data is reported in Study 2). In the first experiment, participants went through a fame- and frequency-orienting task at study in three study/test blocks with 64 names presented per block (of which half were distractors), where participants made old/new responses at test. In the second experiment, 64 names were studied and 128 presented at test with item responses followed by remember/know judgments. In experiment three, participants went through two study/test blocks with 64 studied and 128 test items, where they made item responses followed by confidence responses used to compute ROC curves. In this study, data was collected at two occasions (n=12 and 16), using the same methodological settings. In the fourth experiment, 36 names were studied and 72 names were used as test items in four study/test blocks, where participants made item and confidence responses.

Common for all experiments were the analysis of hits and false alarms. Whereas fame increased hit rates, and lowered false alarms, high frequency lowered hits and increased false alarms, as compared to low frequency (i.e., in correspondence with the mirror effect, see Glanzer & Adams, 1985; 1990a). The name paradigm was also implemented in the remember/know and ROC techniques.

Remember and know responses were not selectively influenced by fame and frequency. Given that the two variables had reliable effects on both remember and know responses, as well as performance (d'), it seems plausible that participants related the two responses to different levels of confidence, rather than to qualities of the encoding experience. Further, the two variables had different effects on response variability, where high fame decreased variability of the old item distribution, whereas lower levels of frequency induced higher old item variability. This effect was not elaborated in the paper (Stenberg, Hellman & Johansson, 2008), but rather served as an argument for the conclusion that fame and frequency affects recognition memory differently.

The results in Study 1 were interpreted as consistent with a dual process account of recognition memory. Fame was associated with increased memory performance as compared to frequency, higher hits and lower false alarms relative frequency and a differential influence on the z-ROC slope in accordance with the DPSD model. It was argued that fame and frequency are associated with two different types of semantic memory, namely specific and non-specific, and that the two types of knowledge relates to recollection and familiarity, respectively.

Study 2 – ERPs of Name Recognition

In the second study, the interaction of episodic and semantic memory was further investigated by using the name paradigm in an ERP recording context. The aim was to replicate the findings from Study 1, and to ascertain the proposed link between fame and recollection, and frequency and familiarity. The study comprised two experiments, both using the name paradigm.

In the first experiment, ERPs were recorded while 288 names were presented divided over 4 study/test blocks where each block contained 36 studied and 72 tested names. EEG was recorded with an electrode cap referenced to the left mastoid and additional electrodes were used to monitor eye movements. Data was digitized at 250 Hz and frequencies at a rate between 0.1 and 30 Hz were accepted.

Each name was presented for 2 sec. during study, where participants tried to remember the name for a subsequent memory test. At test, each name was presented and the testee gave an item response followed by a confidence response (used to compute ROC curves). Behavioral data replicated those in Study 1 (as described in the fourth experiment in that study), and the DPSD model was used to account for ROC data. It was shown that fame and frequency was related to different estimates of recognition, namely recollection and familiarity as defined by the DPSD. ERP difference waveforms were quantified in four intervals: 300-500, 500-700, 700-900 and 900-1100 ms., and it was predicted that frequency would elicit an effect in the earliest interval (the FN400), and fame in the second interval (the LPC). Indeed, frequency resulted in a main effect in the 300-500 ms. interval, whereas fame did not. However, only low frequent names generated the effect. Both fame and frequency elicited a main effect in the 500-700 ms. interval, although the former was larger and the spatial gradients of the two variables were different (the effect of frequency was largest at frontal electrodes and fame induced maximal effects at posterior sites). It was concluded that fame and frequency were differentially related to the LPC and the FN400 old/new effect, supporting the view that recognition memory involves two different retrieval processes.

In the second experiment, twenty-two participants went through a frequency- and celebrity judgment task administered online using the Inquisit software (www.millisecond.com). The aim was to investigate whether familiarity, reflected by the FN400, is confounded with conceptual priming (Paller, Voss & Boehm, 2007), or not. Each participant went through either a frequency or fame judgment task at encoding, and the other at test, where degree of priming was estimated as the difference in RT to re-presented names and new names. Fame had an effect in the task where fame was made salient during measurement of priming, whereas neither fame nor frequency had an effect when the frequency judgment task was implemented at test. That is, frequency was unrelated to conceptual priming, suggesting that the FN400 reflects familiarity.

Study 3 - A formal Account of the name paradigm

To replicate the findings from Study 1 and 2, and to provide a formal account of the effects of pre-experimental familiarity on recognition memory, a novel implementation of the VT (Sikström, 2001) was applied to name memory data. Given this focus in Study 2, little attention was paid to the debate on the nature of recognition memory (i.e., whether a recognition decision is governed by one or two processes). A brief description of the VT is provided here, which is elaborated in Study 2 (Hellman & Sikström, submitted).

The Variance Theory (VT)

The VT describes recognition in terms of the relation between the study material and the pre-experimental context associations induced by the test items. Two separate vectors of binary features, the item and the context layers, represent the to be remembered item and contextual information associated with the item. The contexts and items are represented as binary activation patterns across the nodes in each layer, and a node is activated if the corresponding feature is present, and inactive if the corresponding feature is absent. At encoding, the weights between simultaneously activated nodes changes during encoding according to a Hebbian learning rule (Hopfield, 1982), and the expected value of the net input (i.e., the signal a node receives from other active nodes connected to the particular node) is kept to zero by subtracting the expected probability of active nodes from both vectors. The two types of contexts, the study context and the pre-experimental context, are represented in the network in one common context layer. A key aspect of the model is that the variance of the net inputs to the context nodes (from the item nodes) increases with the number of contexts that are connected to the item, which means that fewer contexts are simultaneously activated for low frequency items than for high frequent ones. At recognition, the item features are reinstated by presentation of an item vector, and, similarly, the features of the study context are reinstated by presentation of the context pattern encoded at the study phase.

Recognition strength is based on the subset of item and context nodes with activation above a specific activation threshold at retrieval. A node is activated at recognition if it was active during encoding and the net input exceeds the activation threshold at retrieval. Whether a “yes” or “no” response is given depends on the subjective recognition criterion, where an item is accepted as old if recognition strength exceeds the criterion.

The analytical solution for the VT is described in Study 2, but some aspects of this implementation should be mentioned here.

New items, which have not been encoded with a context, have a net input equal to the sum of random weights, and because the expected values of all weights are zero, so will the net input be for unstudied items. For old items, where a context has been encoded with the item, the net input is the sum of weights connected to the node whose respective context nodes were active at encoding. Frequency relates to variations in the standard deviation of the net input, and the expected number (and proportion of) active nodes determine the net input. Therefore, an increase in frequency leads to higher variability in net input, which explains why high and low frequent items have different effects on recognition performance. Fame affects the magnitude but not the standard deviation of the net input and relies solely on the expected value of the net input, because fame is implemented as identical context representations. We choose this path because the representation of a famous person is assumed to be less distorted by incremental experiences. We argue that repeated exposure to a famous name strengthens the memory representation of that person because there is a high overlap of context representations during several encoding events. That is, repeated encoding of a famous name does not alter the original representation to the same extent as for a non-famous name, and by implementing fame as identical context representations, the model is constrained with one parameter (in practice, the context representations are highly correlated, but is implemented as identical representations). There is a lower variability in net input for an item with high pre-experimental familiarity and identical contexts, which increases the number of active nodes. For a new item, with a non-encoded active pattern of nodes, the expected value of the net input simply equals the context representation induced by previous encounters.

The VT provides a formal and detailed account of both item frequency and fame, by means of changes in variability of the net input to respective layer for the item. The model is based on two layers in the item representation, using a Hebbian learning rule for the simulation of encoding and defines recognition strength in terms of the proportion of active features in relation to an activation threshold.

To investigate whether the VT could predict performance for fame and frequency, and if the two variables were selectively related to the item and context layers of the model, the VT was applied on two different data sets from a source memory and a paired associates test. Source memory data can be accounted for because the Hebbian learning rule increases the association of item and context information at encoding, and these encoded patterns are reinstated as cues at retrieval. Thus, for items that generate a lower variability of the net input to respective layer and a higher magnitude of the net input, source memory performance increases (i.e., famous items). For associative recognition, two items presented at test evoke the encoded pattern to a certain degree determined by the magnitude of the net input in relation to the activation threshold for active nodes. The item layer represents information about the respective item vector, whereas contextual information about the encoded pattern is reinstated by activation of the context layer. To reactivate information in the context layer related to the encoded pattern and not any previous pattern, both magnitude and variability of the net input to the item and context layers are important.

In the first experiment, 60 students participated and were tested on 96 items (of which half were distractors) divided over two study/test blocks, from a pool of names that was an updated version of the pool used in Study 1. At study, each name was presented in one of two colors, and at test, participants made item responses and source memory responses. In experiment two, 32 participants studied 12 name combinations (presented in one of four spatial locations on the screen – used as the source memory variable) and were tested on 8 old combinations, 8 recombined combinations and 8 new combinations, for which they made item, source and paired associates judgments. This procedure was used in four study/test blocks, with 192 tested names in total.

As predicted by previous research with the name paradigm (Study 1 and 2), and the VT, frequency was inversely related to memory performance, and fame was beneficial for recognition in item, source and paired-associate memory. The results were discussed in the context of single- and dual process theory. The VT is not an explicit dual-process model, albeit the assumption of dual processes in recognition can be accommodated within the VT by relating the context layer to familiarity and both the context and item layer to recollection. In the fit of the model to empirical data, it was shown that average predicted activity in the context and item layers (*mse*) were differently affected by fame and frequency, in accordance with dual-process theory.

Study 4 – Item Variability and Attention

The focus of Study 4 was to implement a new model that provides a comprehensive account of item variability; the MSDT. As described in the theoretical review in the introduction, current recognition memory models lack a detailed description of how the new item distribution affects the latent familiarity distributions, and provides an interpretation of the z-ROC slope limited to item familiarity (SDT) and properties of retrieval processes (the DPSD model). The MSDT is proposed to account for performance, item- and response variability in people with attentional deficits and healthy controls, with a novel suggestion of how to interpret and possibly use the z-ROC measure. The model is shortly described below, and described in detail in Study 4.

The Multidimensional Signal Detection Theory (the MSDT)

The MSDT delineates recognition memory performance, response variability and item variability as the result of activity in N nodes where a varying number of nodes are activated for a presented item. Noise is equally distributed across all nodes (with an expected value of zero and a standard deviation of one) whereas the signal is focused to a subset of (a) number of nodes, where the sum of the signals is a variable (S). A node is active when the content of signal + noise or noise-only exceeds an activation threshold (t).

If (*a*) is high, the signal is distributed to a large number of nodes, resulting in few active nodes with detrimental effects on recognition accuracy. For low values of (*a*), the signal is focused to a few nodes, increasing the number of active nodes. This occurs because the total value of (*S*) is equal and independent of (*a*). The activation threshold can vary in position and for high values of (*t*), few nodes reach the threshold, and at least one active node is demanded for a "yes" response. Consequently, low values of (*t*) results in a higher number of erroneous recognition decisions, that is, false alarms. Thus, the two parameters (*a*) and (*t*) reflects changes in new to old item variability and response bias. Performance is represented by parameter (*S*) – which denotes signal (i.e., corresponds to d'), and interacts with strength variables such as study time and item repetition (i.e., item familiarity). A core feature in the model is that the probability that features are active for old items is larger than the probability that features are active for new items, which also leads to an increase in old item variability.

The MSDT is provided as an account of recognition performance and item variability in attentive and inattentive people, where attentiveness denotes the presence or absence of an attentional deficit (i.e., ADHD or different levels of attentional deficit). The number of nodes receiving signal (*a*, labeled the attention parameter) is larger in inattentive people than attentive, leading to more signal-induced active nodes and lower z-ROC slopes in the latter group. This difference will not be elaborated here, but may be related to dopaminergic function. The model also predicts that inattentive people exhibit a higher overall response variability than attentive people, which represents the common finding of increased response variability in persons with ADHD (Castellanos et al., 2005; Leth-Steensen, Elbaz, & Douglas, 2000). This is modeled by introducing variability in (*t*), which influences performance more in inattentive (large *a*) than attentive people (small *a*).

The MSDT provides a novel and extended account of ROC data for two reasons. First, the model assumes a binomial rather than a normal distribution (as is assumed in the competing models) because recognition depends on the activity of nodes that are either active or non-active. Second, because differences in ROC curves are the result of values of (*a*), the slope of the z-ROC can be related to attention, and this opens up for both a novel interpretation of ROC curves and an interaction of two usually divided research fields; recognition memory and ADHD symptomatology.

Describing ROC data with a binomial distribution has several implications for the understanding of the z-ROC slope, of which some are briefly described here (see Study 3 for a detailed delineation of the ROC account).

First, the z-ROC is commonly interpreted as the ratio of the new to old item standard deviation (Yonelinas & Parks, 2007). When applying this logic to ROC data based on a binomial distribution, the z-slope clearly overestimates the ratio of new to old item standard deviation. Second, the z-ROC measure can be inherently explained by the MSDT because the ratio of new to old item variability is related to the ratio of new to old number of active nodes in the model. Using a binomial distribution instead of a normal, the proportion of active features co-varies with item variability. Third, unequal variability in the latent distribution is the result of a non-linear activation function, and the old item variability exceeds that of the new items due to changes in (*a*). Thus, the MSDT predicts z-ROC slopes below 1.0 in accordance with the literature (Glanzer et al., 1999; Yonelinas & Parks, 2007). However, the MSDT can also account for higher values of the z-ROC slope.

The MSDT predicted that inattentive participants would exhibit higher z-ROC slopes than attentive due to higher values of (*a*) in inattentive people, in contrary to the UVSD, which can be assumed to predict lower z-slopes for inattentive people due to the increased response variability. The DPSD model, on the other hand, has no clear prediction for this topic.

The predictions of the MSDT was tested on participants with high (*n*=45) and low (*n*=30) attention as measured by the 18 item Adult ADHD self-Report Scale Symptom Checklist (ASRS). Participants studied 40 and were tested on 80 concrete nouns in 3 study/test blocks, and at recognition, they made item and confidence responses at test. As expected, inattentive participants had lower performance and higher overall response variability as compared to attentive participants, and also, higher z-ROC slopes. The MSDT was mathematically compared to the UVSD and the DPSD, where empirical data was fitted to a three and four parameter solution. These fittings showed small but existing differences in favor of the MSDT model when the model was fitted to group data, however, when the three models were compared as to MLE and BIC values (Bayesian Information Criterion) when the models were fitted to subject data, there were no significant differences between the models.

Discussion

This dissertation has investigated episodic memory along two paths of inquire; the study of how pre-experimental knowledge affects recognition memory, where the interaction of episodic and semantic memory has been stressed, and the study of processes underlying item variability, where a novel computational model has been introduced and tested.

In the former, investigation of name recognition implies that semantic memory can be both beneficial and detrimental for episodic recognition. It was concluded that these effects are related to different types of semantic memory, and that the proposed differences in conceptual knowledge maps on to different retrieval processes, as suggested by several behavioral experiments where different paradigms and procedures were used (Study 1, 2 and 3), as well as electrophysiological data (Study 2). Further, as outlined in Study 3, the differential effects of fame and frequency on recognition memory can also be understood as alterations in the interaction of item and context information about the test item. These alterations are induced by frequency of occurrence and the degree of similarity of the memory representation at different encoding events.

In the second path of inquire, several aspects of recognition memory and item variability has been investigated. The MSDT provides a new understanding of item variability where a parsimonious 3-parameter model accounts for recognition memory, ROC data and response variability, and links recognition memory to the study of attentional deficits. The main finding is that the conventional interpretation of the slope of the z-ROC as an assessment of new to old item variability and form of the latent distribution can be extended to an estimate of differences in attention.

Prior Knowledge and Recognition Memory

The first study (Stenberg, Hellman, & Johansson, 2008) took an interest in how variations in frequency and fame affect episodic memory given the contradictory ideas of how prior experience influences memory. On the one hand, episodic memory is dependent on semantic memory when an episode is encoded, because general knowledge act as a framework involved in the creation of a new episodic memory (Bartlett, 1932). On the other hand, it is a ubiquitous finding that high frequency impairs memory in comparison with memory for rare items (the word frequency effect), where the former increases false alarms and lowers hits, whereas the latter has the opposite effect, a pattern known as the mirror-effect (Glanzer & Adams, 1990a). The effect of word frequency, and particularly the mirror effect, has been a subject of thorough investigation (Dennis & Humphreys, 2001; Glanzer & Adams, 1990a; Glanzer et al., 1993; Malmberg et al., 2002; Malmberg & Murnane, 2002;McClelland & Chappel, 1998; Shiffrin & Steyvers, 1997). Different accounts have been provided for the mirror effect, where frequency has been given a certain role; as a predictor of the number of context associated with an item (Dennis & Humphreys, 2001; Sikström, 2001), as correlated with commonness of orthographic features and normative versus letter frequency (Malmberg et al., 2002; Malmberg & Murnane, 2002) and as a mediator of changes in variability in a connectionist memory system related to high and low frequency (Sikström, 2001). Others have argued that high and low frequency have differential affects on endorsement of old and new items because the stimulus classes induce differences in attention (Glanzer & Adams, 1990a), in encoding variability (McClelland & Chappell, 1998), and in the selective influence of familiarity and recollection (Joordens & Hockley, 2000). Even others have argued that the word frequency effects is misleading, because word frequency is confounded with lexicality, or orthographic similarity (Estes & Maddox, 2002). That is, the mirror effect is driven by orthographic word likeness rather than item frequency because the authors found no mirror effect when pre-familiarization for the test items was used. Study 1 reported effects on hits and false alarms that were consistent with the mirror effect, both for frequency and fame. For variations in frequency, the effect can be related to item frequency because high frequency lead to an increase in hits and a decrease in false alarms relative the low frequency names. However, also variations in fame resulted in a mirror effect, which arguable is not related to orthographic similarity because a famous name, may it be rare or common in terms of frequency, is associated with a certain degree of specificity in the representation of the famous individual that seems to induce the name frequency effect. Word frequency effects can be related to other qualities than frequency of occurrence, but name frequency evokes a memory representation encompassing information about a certain individual. For common, non-famous names, this information is relatively unrelated to individual characteristics, but for famous names, the retrieved information is highly autobiographical, independent of item frequency. Also, manipulation of pre-familiarization, which should reduce the mirror effect according to Estes and Maddox (2002), shares similarities with manipulation of fame because famous names evoke highly similar context representations at different encoding events, as in in Study 1, 2 and 3. Still, the mirror effect emerged.

Yet another possible explanation for the name - mirror effect is that is occurs as the result of metacognitive diagnostic derivation. That is, if the participant realizes that a famous name is easier to remember than a non-famous one, the participant would simply reject an unstudied famous name because had it been studied, it would remembered. However, since this metacognitive memory dimension was not manipulated in any of the experiments in Study 1, 2 or 3, it is not possible to validate the plausibility of the hypothesis with the present data. However, Palmer (2007) investigated the influence of metamemory on the mirror effect in six experiments, where different experimental manipulations that tax metamemory processing were used. The study revealed that hits were affected by metamemory processing, whereas the false alarm rate remained unaffected by these manipulations.

In Study 1 and 2, the difference in memory performance and item variability was related to different retrieval processes. It was argued that high frequency impaired memory because a name that is unrelated to a specific individual omits deep encoding as it lacks novelty, and thereby modulates encoding and retrieval processes negatively. Fernandez and Tendolkar (2006) suggested that the rhinal cortex, a cortical region closely associated with the hippocampus, acts as a novelty detector. Correspondingly, a novel item will gain access to encoding resources as the result of processing in additional structures of the medial temporal lobe, whereas a familiar item is omitted such memory-facilitating processing. At retrieval, the item will induce a sense of familiarity as a result of previous encoding, albeit lacking a reliable link between the probe and the item in memory, leading to increases in both false alarms, z-ROC slope and induce an old/new effect different from that in fame recognition. A rare name, on the other hand, will benefit encoding resources because it is novel, however, at retrieval, recognition of a low frequent name (relative a famous name) suffers because there is no individuating features related to the name. A related perspective was offered by Mandler (1980; 2008), stressing the importance of an increment in familiarity as a function of previous occurrence. A common name has been encountered in a multitude of episodes and contexts, meaning that an additional encoding leads to a relatively low increase in memory strength. A rare name will benefit from the lack of previous encounter because the increment in memory strength caused by an additional encoding episode is relatively high given the low degree of previous integration. Thus, it is the familiarity of the item that leads to higher performance for rare than common names. The hypothesis brought forward by Mandler was supported by results in Study 2, where only low frequent names elicited the electrophysiological signature of familiarity, the FN400, whereas frequent names did not.

In Study 3, the detrimental effect of frequency on recognition memory was understood as a modulation of the variability in the context representation of the item. That is, because a frequent name has been encoded in several previous events, it is associated with a high number of pre-experimental contexts. The increase in pre-experimental contexts raises the variability of the input, which has a negative net effect of the number of active nodes, and thereby, on familiarity. Fame, on the other hand, was assumed to induce pre-experimental contexts similar to those evoked during experimental study, and that a famous name is encoded with a higher specificity. Thus, knowledge about a famous person is based on encoding of relevant autobiographical information, which can also induce affective responses. Therefore, the representation of a famous person is more stable than that of a non-famous name. Fame was therefore implemented as identical pre-experimental context representations in the VT.

It can be argued that the high specificity of the representation of a famous name in memory (Stenberg et al., 2009; Zion-Golumbic et al., 2010) is related to retrieval of item and contextual information with high similarity over different encoding events. Thus, fame recognition can be accommodated within the framework of the encoding – specificity principle, or the ESP (Tulving & Thompson, 1973). According to the ESP, memory performance is to a high degree determined by the consistency of contextual information at encoding and retrieval. The ESP has accounted for changes in memory performance for several manipulations, such as difference in strength of retrieval cues (Adam et al., 2007), variations in physical environment (Grant et al., 1998), differences in auditory environment (Godden & Baddely, 1980) and varying degree of intoxication (Weingartner et al., 1976). Thus, the specificity of encoding operations, which determines what will be stored in memory and thereby which retrieval cues will be effective for subsequent retrieval, is of importance for memory performance. Even though the ESP is conventionally studied with manipulation of exogenous variables as described above, the principle can be applied to differences in contextual information for a specific memory item. For example, it has been shown that memory performance benefit from semantically related cues present at encoding and retrieval (Reder, Anderson, & Bjork, 1974). This implies that contextual information about an item (i.e., information about a famous person, such as movie character, voice, etc.) that does not vary over encoding events (the memory representation of Javier Bardem in the motion picture “No country for old men” is unaffected by subsequent encoding of the name Javier Bardem) serves as an efficient retrieval cue.

The effects of frequency on recognition memory, as discussed above, maps on to a certain type of semantic memory because knowledge about frequency of occurrence modulates recognition of high and low frequent items. This non-specific type of semantic memory provides a vague sense of familiarity for the presented item at test, with no access to individuating features of the item necessary to involve the name in the fabric of our own lives (i.e., episodic memory). A famous name, on the other hand, has a richness of information related to the name, which is strongly associated with the stored name in memory. At encoding, individuating (episodic) information about the name is retrieved and stored with the item, in addition to distinctive semantic information. At test, the probe induces successful retrieval of item and context information because the accessibility to the item relies on several retrieval cues (i.e., individuating features of the name) that match the stored representation of the item to a high degree, and because there is a high similarity between the probe and information evoked by previous encoding episodes. Thus, recognition of a famous name benefits from the distinctiveness of the contextual information about the item, the strong link between contextual and item information and the amount of contextual information acting as retrieval cues at test. Further, because previous experience of a famous name has resulted in an episodic memory of the name, retrieval of contextual information occurs (i.e., recollection). Because individuating features links the presented item to previous experience, and the highly specific associations makes the memory more distinctive at encoding, recognition accuracy increases, which results in higher hit rates with no or small effects on false alarms.

The psychological implications of the result of Study 1 and 2 described above provide a sufficient understanding of fame and frequency effects. However, the interpretations in these studies may seem contradictory to that in Study 3 in one respect, namely regarding the issue of whether recognition memory is governed by one or two memory processes. The results in Study 1 and 2 are interpreted as that specific and non-specific semantic memory induces recognition by recollection and familiarity, respectively. In Study 3, fame and frequency was related to partially dissociated activity in the item and context layers, respectively. Following the common interpretation of results from associative and source memory paradigms, both experiment 1 and 2 in Study 3 supported the conclusion in Study 1 and 2: that two processes contribute to recognition with differential effects on item, source and associative memory. This was accommodated in the VT by dividing the influence of fame and frequency on the standard deviation in net input and the magnitude in net input to the two layers (item and context). Thus, we argue that the presumed selective influence of familiarity and recollection can be described by the VT, where fame and frequency have different effects on activity in the context and item layers. According to the averages for item and context layer activity in Experiment 1 and 2 in Study 2, for old items, it seems that fame increased the activity in both the item and the context layer relative non-famous names. Frequency induced no difference in the item layer but high fame lead to lower activity in the context layer. This pattern also emerged in the fit to item and source memory data in Study 2. In the fit to associative memory, a task were recollection is more salient, there were no difference for high and low frequency in neither the item nor the context layer, but a difference in averages in the item layer for fame.

Given the behavioral and electrophysiological findings from Study 1, 2 and 3, it is concluded that retrieval of contextual information, which is necessary for successful source- and associative memory, is partially dissociated from item recognition, where additional episodic information is not necessary for a correct memory decision. That is, item memory (i.e., familiarity) occurs either when additional information about the item cannot be retrieved, or when such additional retrieval is unnecessary for the task. Study 3 revealed that this recognition difference occurs because some items (non-famous, high frequent names), more than others (famous names), induce an increased variability in input in the network where information about an item and its experimental and pre-experimental contexts are processed. This variability is related to the degree and type of pre-experimental encodings, that is, if a high number of pre-experimental encodings is associated with a high or low degree of consistency of previous and current context representations (non-famous vs. famous names). This means that the seemingly contradictory views on recognition memory in Study 1 and 2, and Study 3, rather reflects different aspects of the recognition decision process. An item that is associated with a high number of pre-experimental encodings can either elevate or impair recognition memory, depending on whether information from the pre-experimental encodings (the knowledge about the item) are consistent with each other regarding semantic content, or are spread and not specifically related to the item at hand. This variability in processing of contextual information leads to recognition with or without access to episodic information that in turn results in a complete instantiation of the study event (recollection), or to a vague familiarity with the item at hand.

A Novel Interpretation of the z-ROC Slope

In Study 4, focus was put on item variability, or confidence data. By recording Receiver-Operating Characteristics (ROCs), it is possible to study memory while controlling for bias and performance. This paradigm has been used thoroughly in the recognition memory literature to understand the underpinnings of memory decisions and to develop and compare models of memory. In the conventional use of ROC data, two measures have been of certain interest, namely the slope and the shape of the z-ROC. The former is assumed to reflect the level of variability in the old item distribution relative the new item distribution, whereas the latter implies whether the latent distributions are Gaussian or not. The latter is important because the form of the distribution yields information about the processes underlying recognition memory (for more details, see Yonelinas & Parks, 2007). In the present thesis, the slope of the z-ROC was stressed, but the MSDT also revealed an interesting point regarding the shape of the z-ROC. In essence, the MSDT suggests that the conventional interpretation of the z-slope may be insufficient, and should be extended, and that recognition memory may be better understood if the assumption of a Gaussian latent distribution is reconsidered. These questions will be discussed below.

In the majority of recognition memory models, a normal underlying distribution is assumed. That this is the common assumption is not hard to understand, because a normal distribution has several advantages. First, it is tractable and analytical because it is easy to derive the probability content within a certain number of standard deviations. Second, the bell-curved distribution is familiar and adequate to describe a multitude of behaviors. Third, given the central limit theorem (i.e., if a sufficient large number of independent random variables with a defined mean and variance exists, the resulting distribution will be Gaussian), a normal distribution can be used to approximate different kinds of distributions in large samples. However, to use the normal distribution in a model of recognition memory, one must assume that the described behavior is continuous (e.g., familiarity). Doing so may lead to exclusion of certain details, or aspects of the underlying layer, or dimension of the described behavior (i.e., neurocognitive activity etc.). In the MSDT, the binomial distribution is used because the model describes an earlier stage of activity than familiarity. That is, the MSDT assumes that neural activity can be represented by node activity, and because these nodes are either active (e.g., post-synaptic activity) or inactive (e.g., no presynaptic summation leading to action potential), the binomial distribution is adequate. To use this distribution rather than the Gaussian is also reasonable because a recognition memory decision should be described by a random binomial variable – either you remember, or you do not, and each memory decision is independent of the former or latter ones. However, can usage of the binomial distribution live up to the advantages of the Gaussian distribution? It turns out it is not only sufficient to describe recognition memory, it also provides novel insights of item variability.

The binomial is as tractable and analytical as the Gaussian, with a defined mean and variance, and because the described variable is discrete, the delineated behavior does not omit any information in comparison with the case of the normal distribution. An advantage of the binomial is that it is naturally skewed. For low values of p and n, say 15 independent trials and a probability of success of 0.2, the distribution is skewed right, meaning that the majority of the probability falls within low numbers of the random variable (X =1, 2 and 3). For higher numbers of both n and p, the distribution gets more symmetrical (i.e., Gaussian). This is adequate because, as described above, the number of expected active nodes for successful recognition is low. Further, and of more interest in the MSDT, the variance is not constant, but is determined by n and p. This means that for high values of p, the variance is low, and when p = 0.5, the variance is at maximum, and there after the variance decreases. The assumption that the probability for active nodes is low seems plausible because items are sparsely represented in the brain, and because the probability of active nodes is larger for old than new items. Thus, the different variability for old and new items derived by the MSDT is the result of the features of the underlying distribution.

The MSDT was used in Study 4, where it described recognition memory in attentive and inattentive people. Differences in performance, response variability and item variability were described with variations in two parameters, the attention parameter (i.e., the number of nodes with input of signal) and the activation threshold. Because the activation threshold is affected by attentiveness, inattentive people have a higher variability in (*t*), leading to lower performance, more specifically, higher numbers of false alarms. This was confirmed by empirical data. The difference in item variability was, as described above, related to the allocation of signal (*a*). However, Study 4 indicated that the difference in z-ROC slope was in fact a result of changes in the new item distribution, rather than the old item distribution. This is interesting because both SDT and the DPSD model assumes that the z-slope is modulated by changes in the old item distribution (i.e., due to encoding variability or the impact of recollection), whereas item variability is more or less uninfluenced by new items. The MSDT suggests the contrary, and it is proposed that SDT and the DPSD model do not stress the importance of prior knowledge enough. When a set of words is studied and the task is to distinguish studied from unstudied items, it seems reasonable that the processes underpinning the recognition decision is modulated by the familiarity of the presented word. Thus, when presented with a pre-experimentally familiar word that has been omitted a deepened encoding - we are likely to endorse it as old due to the familiarity of the item.

As described in Study 4, and in this discussion, the MSDT both highlight limitations with the current framework of item variability (SDT and the DPSD model), and provides novel insights for the study of recognition memory and ROC data for attentive and inattentive people. The main focus has been the z-ROC, and the MSDT suggest that the study of confidence data and ADHD symptomology, two fields that are commonly separated, can in combination gain novel knowledge from each other. Further, the model suggests that both differences in memory performance, response variability and item variability over different levels of attention are mediated by variations in density of node activity. This is not analogous to a specific level of neural activity, but indicates that the relative number of operating features at the neural level is important for the described effects.

Limitations

The present thesis includes a variety of experimental paradigms, methods and techniques and different modeling procedures. However, to understand what the result of these experiment and modeling attempts can reveal about the concepts of interest, it is important to acknowledge the scientific limitations that follow these procedures. Below, the most obvious caveats and possible criticism with the present studies are discussed.

A possible critique for the three first studies, where the name paradigm was used, is that the effects brought by the frequency manipulation may in fact be mediated by another, possibly lexical, confounding variable. This problem was described in the discussion, were it was argued that it is unreasonable that the proposed name frequency effect (that occurred for both fame and frequency) is the result of some kind of lexical variable (i.e., orthographic similarity), because names and words differ substantially in terms of the experience of orthographic irregularities and processing. The latter refers to how the representation of a name is different from that of a word. A noun, may it be common or rare, evokes a certain amount of episodic and semantic associations stemming from the object denoted by the word, say “closet” or “carboy”. Even though associated conceptual information (“clothes” or “gallows”, and “wine” or “grapes”) and biographical associations (re-experiencing this morning when I got dressed, or the first time I saw a carboy) may be more or less vivid and specific, it is different from processing of a name. Tom Cruise brings me back to one or several cinematic characters, a face and a voice, remembrance of interviews, associations of a friend who really loves (or dislikes) the actor, and so on. This is much and vivid information that initializes a series of associations, thereby increasing the distance between the letters forming the name and the induced content (the associations). Therefore, it is hard to see how word or feature similarity would influence remembrance a famous name. The same logic applies to non-famous names. The orthographic irregularity of the name Rosamund Dankworth (yes, some people bear that name) brings few associations to mind, which may put more demand on shallow (lexical) processing of the item. However, comparing the processing of “carboy”, or “crampon”, and Rosamund Dankworth differs because the former more often brings an object to mind, thus evoking a more specific representation. Even though common, non-famous names may seem more similar to common words in this respect, it is a personal name related to a more or less known individual. Of cause, it is possible that similarity of common names affects encoding operations and leads to worse performance at test, but an answer to that question demands additional investigation.

It is argued that the VT can describe empirical data by assuming dual-processes in recognition memory by relating activity in the item and context layers selectively to fame and frequency, and that the two variables corresponds to recollection and familiarity, respectively. Even though this conclusion is theoretically plausible, is has not been verified by empirical data, because the predicted activity in the item and context layers are averages generated during the fit to group data. Thus, the observed differences in item and context layer activity related to fame and frequency cannot be tested for significance, and therefore only suggests that a partial dissociation of the influence of the experimental variables on activity in the models layers may be possible. It is therefore important to investigate this further in future research, by (for instance) using the VT to account for empirical data where different paradigms commonly used to estimate the influence of familiarity and recollection are implemented.

Further, because the VT is not compared with any other memory model, it is not possible to infer how well the model describes empirical data and the construct of interest. Results of the achieved model fit can be interpreted as if the model provides a reasonable good account of data, because the predicted response probabilities are encompassed by the confidence interval for the sample responses. However, to ascertain that the model accounts for fame and frequency effects in recognition memory, the fit of the model should be compared to, for instance, a model fit of the UVSD and the DPSD models.

In Study 4, the MSDT provides an account and novel insights on recognition memory data and item variability over different levels of attention. However, the inattentive group is also referred to as people with ADHD. The means for estimating the level of attentiveness is a standard questionnaire commonly used to assert whether a participant has ADHD related attentional deficits, which opens up for the possibility that some testees have made (more or less deliberate) incorrect responses. This may introduce an uncertainty to what qualities the analysis, based on the division of data over levels of attentiveness, really constitute.

There are also arguments in Study 4 that rests on seemingly perilous assumptions. It is argued that the z-ROC slope can be an indicative of dopamine levels. This is based on the connection between the z-slope and the attention parameter (*a*), and because low levels of (*a*) is thought to reflect attentional deficit where dopamine is known to be involved, the z-slope is related to dopamine because (*a*) modulates the z-ROC. Thus, there is a rather long and complex causative chain, and because no biological data are collected, the assumption cannot be empirically tested with the present data. Now, the link between attention (ADHD) and dopamine is thoroughly studied, and therefore introduces no particular reason to question the basic assumption. However, and as discussed above, it may be perilous to assume that the recorded data is really drawn from a sample of people suffering from ADHD. Further, to verify that the z-ROC slope is sensitive to changes in dopaminergic state, rather than differences in performance and response bias, demands additional studies, preferably using the rat model where dopamine can be controlled and z-ROCs can be collected in an odor-learning paradigm similar to that used by Fortin and colleagues (Fortin, Wright, & Eichenbaum, 2004). The reason performance may be a confounding variable is that accuracy and z-ROC slope co-varies, as do (*a*) and the activation threshold (response bias). It is therefore important to differentiate the effects of these parameters to elucidate whether the z-slope can be used as a measure of dopamine levels.

Another limitation with the model fit in Study 4 is the fact that the mean MLE was used to compare the different models and implementations. Had the aggregated MLE been computed in the fitting to both group- and individual data, a direct comparison could have been made between the item and group fits. To use the aggregated MLE for this purpose would be necessary because the number of underlying, logarithmic probabilities differ in mean MLE for group and individual fit, which is not the case when aggregated MLE is used.

References

Adam, S. M., Van der Linden, A., Ivanoiu, A.-C., & Juillerat, S., Bechet, E. S. (2007). Optimization of encoding specificity for the diagnosis of early AD: The RI-48 task. *Journal of Clinical and Experimental neuropsychology 29*(5), 477–487.

Addante, R. J., Ranganath, C., Olichney, J., Yonelinas, A. P. (2012). Neurophysiological evidence for a recollection impairment in amnesia patients that leaves familiarity intact, *Neuropsychologia, 50* (13), 3004-3014.

Aggleton, J. P., & Brown, M. W. (2006). Interleaving brain systems for episodic and recognition memory. *Trends Cogn Sci, 10*(10), 455-463.

Aly, M., Knight, R. T., & Yonelinas, A. P. (2010). Faces are special but not too special: spared face recognition in amnesia is based on familiarity. *Neuropsychologia, 48*(13), 3941-3948.

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychol Sci, 2*(6), 396-408.

Arndt, J., & Reder, L. M. (2002). Word frequency and receiver operating characteristic curves in recognition memory: evidence for a dual-process interpretation. *J Exp Psychol Learn Mem Cogn, 28*(5), 830-842.

Badre, D., & Wagner, A. D. (2002). Semantic retrieval, mnemonic control, and prefrontal cortex. *Behav Cogn Neurosci Rev, 1*(3), 206-218.

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

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychon Bull Rev, 6*(1), 57-86.

Bayley, P. J., Hopkins, R. O., & Squire, L. R. (2006). The fate of old memories after medial temporal lobe damage. *J Neurosci, 26*(51), 13311-13317.

Berry, C. J., Shanks, D. R., & Henson, R. N. (2008). A unitary signal-detection model of implicit and explicit memory. *Trends Cogn Sci, 12*(10), 367-373.

Castellanos, F.X., Sonuga-Barke, E. J. S., Scheres, A., Martino, A.D., Hyde, C., & Walters, J. R. (2005). Varieties of attention-deficit/hyperactivity disorder-related intra-individual variability. *Biol Psychol, 57*, 1416-1423.

Chalmers, K. A., & Humphreys, M. S. (1998). Role of generalized and episode specific memories in the word frequency effect in recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 24*(3), 610-632.

Clark, S. E., & Gronlund, S. D. (1996). Global matching models of recognition memory: How do the models match the data. *Psychon Bull Rev, 3*, 37-60.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(671-684).

Curran, T. (2000). Brain potentials of recollection and familiarity. *Mem Cognit, 28*(6), 923-938.

Curran T. (1999). The electrophysiology of incidental and intentional retrieval: ERP-old/new effects in lexical decision and recognition memory. *Neuropsychologia, 37*, 771–785.

Curran, T. (2004). Effects of attention and confidence on the hypothesized ERP correlates of recollection and familiarity. *Neuropsychologia, 42*(8), 1088-1106.

Dennis, S. (1995). The impact of the environment on the word frequency and the null list strength effects in recognition memory [Electronic version]. Noetica, 1. Retrieved 19

March 2007 from http://www2.psy.uq.edu.au/CogPsych/Noetica/Issue1/Environment.html
Dennis, S., & Humphreys, M. S. (2001). A context noise model of episodic word recognition. *Psychol Rev, 108*(2), 452-478.
Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2007). Imaging recollection and familiarity in the medial temporal lobe: a three-component model. *Trends Cogn Sci, 11*(9), 379-386.
Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2008). The effects of unitization on familiarity-based source memory: testing a behavioral prediction derived form neuroimaging data. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(4), 730-740.
Duarte, A., Ranganath, C., Winward, L., Hayward, D., & Knight, R. T. (2004). Dissociable neural correlates for familiarity and recollection during the encoding and retrieval of pictures. *Cogntive Brain Research, 18*, 255-272
Dunn, J. C. (2004). Remember-know: a matter of confidence. *Psychol Rev, 111*(2), 524-542.
Dunn, J. C. (2008). The dimensionality of the remember-know task: a state-trace analysis. *Psychol Rev, 115*(2), 426-446.
Ebbinghaus, H. (1885). Über das gedächtnis: Untersuchungen zur experiementellen psycholgie. *Leipzig: Duncker & Humbolt*.
Egan, J. P. (1958). Recognition memory and the operating characteristic. Technical note AFCRC-TN-58-51, Hearing and communication laboratory, Indiana University.
Estes, W. K., & Maddox, W. T. (2002). On the processes underlying stimulus-familiarity effects in recognition of words and nonwords. *J Exp Psychol Learn Mem Cogn, 28*(6), 1003-1018.
Fernandez, G., & Tendolkar, I. (2006). The rhinal cortex: "Gatekeeper" of the declarative memory system. *Trends in Cognitive Sciences, 10(8)*, 358-362.
Fortin, N. J., Wright, S. P., & Eichenbaum, H. (2004). Recollection-like memory retrieval in rats is dependent on the hippocampus. *Nature, 431*(7005), 188-191.
Friedman, D., Johnson, Jr. R. (2000). Event-related potential (ERP) studies of memory encoding and retrieval: a selective review. *Microscopy Research and Technique, 51*, 6–28.
Glanzer, M., & Adams, J. K. (1985). The mirror effect in recognition memory. *Mem Cognit, 13*, 8-20.
Glanzer, M., & Adams, J. K. (1990a). The mirror effect in recognition memory: Data and theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(1), 5-16.
Glanzer, M., & Adams, J. K. (1990b). The mirror effect in recognition memory: data and theory. *J Exp Psychol Learn Mem Cogn, 16*(1), 5-16.
Glanzer, M., Adams, J. K., Iverson, G. J., & Kim, K. (1993). The regularities of recognition memory. *Psychol Rev, 100*(3), 546-567.
Glanzer, M., Hilford, A., & Kim, K. (2004). Six regularities of source recognition. *J Exp Psychol Learn Mem Cogn, 30*(6), 1176-1195.
Glanzer, M., Kim, K., Hilford, A., & Adams, J. K. (1999). Slope of the receiver-operating characteristic in recognition memory. *Journal of Experimental Psychology-Learning Memory and Cognition, 25*(2), 500-513.
Graham, K. S., Becker, J. T., & Hodges, J. R. (1997). On the relationship between knowledge and memory for pictures: evidence from the study of patients with semantic dementia and Alzheimer's disease. *J Int Neuropsychol Soc, 3*(6), 534-544.

Graham, K. S., Simons, J. S., Pratt, K. H., Patterson, K., & Hodges, J. R. (2000). Insights from semantic dementia on the relationship between episodic and semantic memory. *Neuropsychologia, 38*(3), 313-324.

Godden, D. R., Baddeley, A. D. (1980). When does context influence recognition memory?. *British Journal of Psychology, 71*, 99–104.

Grant, H. M., Bredahl, L. C., Clay, J., Ferrie, J., Groves, J. E., McDorman, T. A., Dark, V. J. (1998). Context-dependent memory for meaningful material: Information for students. *Applied Cognitive Psychology* , *12*, 617–623.

Greenberg, D. L., & Verfaellie, M. (2010). Interdependence of episodic and semantic memory: evidence from neuropsychology. *J Int Neuropsychol Soc, 16*(5), 748-753.

Gronlund, S. D., & Elam, L. E. (1994). List-length effect: Recognition accuracy and variance of underlying distributions. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 20*(6), 1355-1369.

Habib, R., Nyberg, L., & Tulving, E. (2003). Hemispheric asymmetries of memory: the HERA model revisited. *Trends Cogn Sci, 7*(6), 241-245.

Healy, M. R., Light, L. L., & Chung, C. (2005). Dual-process models of associative recognition in young and older adults: evidence from receiver operating characteristics. *J Exp Psychol Learn Mem Cogn, 31*(4), 768-788.

Heathcote, A. (2003). Item recognition memory and the receiver operating characteristic. *J Exp Psychol Learn Mem Cogn, 29*(6), 1210-1230.

Hintzman, D. L. (1988). Judgment of frequency and recognition memory in a multiple trace memory model. *Psychol Rev, 95*, 528-551.

Hintzman, D. L. (1994). On explaining the mirror effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(1), 201-205.

Hirshman, E., & Hostetter, M. (2000). Using ROC curves to test models of recognition memory: the relationship between presentation duration and slope. *Mem Cognit, 28*(2), 161-166.

Hodges, J.R., & Patterson, K. (1995). Is semantic memory consistently impaired early in the course of Alzheimer’s disease? Neuroanatomical and diagnostic implications. *Neuropsychologia, 33*, 441–459.

Hopfield, J. J. (1982). Neural Networks and Physical Systems with Emergent Computational Abilities. *Proceeding of the National Academy of Sciences, USA, 81*, 3088-3092.

Howard, M. W., Bessette-Symons, B., Zhang, Y., & Hoyer, W. J. (2006). Aging selectively impairs recollection in recognition memory for pictures: evidence from modeling and receiver operating characteristic curves. *Psychol Aging, 21*(1), 96-106.

Jang, Y., Mickes, L., & Wixted, J. T. (2012). Three tests and three corrections: comment on Koen and Yonelinas (2010). *J Exp Psychol Learn Mem Cogn, 38*(2), 513-523.

Johnson, M. K. (2006). Memory and reality. *Am Psychol, 61*(8), 760-771.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychol Bull, 114*(1), 3-28.

Joordens, S., & Hockley, W. E. (2000). Recollection and familiarity though the looking glass: When old does not mirror new. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(6), 1534-1555.

Kinchla, R. A. (1994). Comments on Batchelder and Riefer 's Multinomial model for source monitoring. *Psychol Rev, 101*(1), 166-171.

Kinsbourne, M., Rufo, D. T., Gamzu, E., Palmer, R. L., & Berliner, A. K. (1991). Neuropsychological deficits in adults with dyslexia. *Dev Med Child Neurol, 33*(9), 763-775.

Kitchener, E. G., Hodges, J. R., & McCarthy, R. (1998). Acquisition of post-morbid vocabulary and semantic facts in the absence of episodic memory. *Brain, 121 (Pt 7)*, 1313-1327.

Koen, J. D., & Yonelinas, A. P. (2010). Memory variability is due to the contribution of recollection and familiarity, not to encoding variability. *J Exp Psychol Learn Mem Cogn, 36*(6), 1536-1542.

Koen, J. D., & Yonelinas, A. P. (2013). Still no evidence for the encoding variability hypothesis: a reply to Jang, Mickes, and Wixted (2012) and Starns, Rotello, and Ratcliff (2012). *J Exp Psychol Learn Mem Cogn, 39*(1), 304-312.

Lepage, M., Ghaffar, O., Nyberg, L., & Tulving, E. (2000). Prefrontal cortex and episodic memory retrieval mode. *Proc Natl Acad Sci U S A, 97*(1), 506-511.

Leth-Steensen, C., Elbaz, Z. K., & Douglas, V.I. (2000). Mean response times, variability, and skew in the responding of ADHD children: a response time distributional approach. *Acta Psychol (Amst), 104*, 167-190.

Luck, S. J. (2005) *An introduction to the Event-Related Potential technique*. The MIT Press, Cambridge.

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

Maddox, W. T., & Estes, W. K. (1997). Direct and indirect stimulus-frequency effects in recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 23*, 539-559.

Malmberg, K. J. (2008). Recognition memory: a review of the critical findings and an integrated theory for relating them. *Cogn Psychol, 57*(4), 335-384.

Malmberg, K. J., & Nelson, T. O. (2003). The word frequency effect for recognition memory and the elevated-attention hypothesis. *Mem Cognit, 31*(1), 35-43.

Malmberg, K. J., Steyvers, M., Stephens, J. D., & Shiffrin, R. M. (2002). Feature frequency effects in recognition memory. *Mem Cognit, 30*(4), 607-613.

Malmberg, K., & Murnane, K. (2002). List composition and the word-frequency effect for recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 28*(4), 616-630.

Mandler, G. (1980). Recognizing: The Judgment of previous occurrence. *Psychol Rev, 87*, 252-271.

Mandler, G. (2008). Familiarity breeds attempts: A critical review of dual process theories of recognition. *Perspectives in Psychological Science 3 (5)*, 392–401.

McClelland, J. L., & Chappell, M. (1998). Familiarity breeds differentiation: A subjective-likelihood approach to the effects of experience in recognition memory. *Psychol Rev, 105*(4), 724-760.

McKoon, G., & Ratcliff, R. (1986). Inferences about predictable events. *J Exp Psychol Learn Mem Cogn, 12*(1), 82-91.

Mecklinger A. (2000). Interfacing mind and brain: a neurocognitive model of recognition memory. *Psychophysiology, 37*, 565–582.

Mecklinger, A. (2006). Electrophysiological measures of familiarity memory. *Clinical EEG and Neuroscience, 37*, 292-297

Mecklinger, A., Meinshausen, R. M. (1998). Recognition memory for object form and object location: an event-related potential study. *Memory and Cognition, 26*, 1068–1088.
Mickes, L., Johnson, E. M., & Wixted, J. T. (2010). Continuous recollection versus unitized familiarity in associative recognition. *J Exp Psychol Learn Mem Cogn, 36*(4), 843-863.
Mickes, L., Wais, P. E., & Wixted, J. T. (2009). Recollection is a continuous process: implications for dual-process theories of recognition memory. *Psychol Sci, 20*(4), 509-515.
Mickes, L., Wixted, J. T., & Wais, P. E. (2007). A direct test of the unequal-variance signal detection model of recognition memory. *Psychon Bull Rev, 14*(5), 858-865.
Murdock, B. B. (1974). *Human memory: Theory and data*: Potamca, MD: Erlbaum.
Neville, H. J, Kutas, M., Chesney, G., Schmidt, A. L. (1986). Event-related brain potentials during initial encoding and recognition memory of congruous and incongruous words. *Journal of Memory and Language, 25*, 75–92.
Nyberg, L., Tulving, E., Habib, R., Nilsson, L. G., Kapur, S., Houle, S., . . . McIntosh, A. R. (1995). Functional brain maps of retrieval mode and recovery of episodic information. *Neuroreport, 7*(1), 249-252.
Paller, K. A., Kutas, M. (1992). Brain potentials during memory retrieval provide neurophysiological support for the distinction between conscious recollection and priming. *Journal of Cognitive Neuroscience, 4*, 375–391.
Paller, K. A., Voss, J. L., & Boehm, S. G. (2007). Validating neural correlates of familiarity. *Trends in Cognitive Sciences, 11*, 243–250.
Paller, K. A., Lucas, H. D., & Voss, J. L. (2012). Assuming too much from 'familiar' brain potentials. TICS, 16(6), 313-315.
Palmer, J. G. (2007). Metamemory and the word frequency mirror effect. Master of Arts thesis, Simon Fraser University. Downloaded from: http://summit.sfu.ca/item/8346, 20/7 2013.
Qin, S., Rijpkema, M., Tendolkar, I., Piekema, C., Hermans, E. J., Binder, M., . . . Fernandez, G. (2009). Dissecting medial temporal lobe contributions to item and associative memory formation. *Neuroimage, 46*(3), 874-881.
Ratcliff, R., McKoon, G., & Tindall, M. (1994). Empirical generality of data from recognition memory receiver-operating characteristic functions and implications for the global memory models. *J Exp Psychol Learn Mem Cogn, 20*(4), 763-785.
Ratcliff, R., Sheu, C. F., & Gronlund, S. D. (1992a). Testing global memory models using ROC curves. *Psychol Rev, 99*(3), 518-535.
Ratcliff, R., Sheu, C.-F., & Gronlund, S. D. (1992b). Testing global memory models using ROC curves. *Psychol Rev, 99*, 518-535.
Ratcliff, R., Van Zandt, T., & McKoon, G. (1995). Process dissociation, single-process theories, and recognition memory. *J Exp Psychol Gen, 124*(4), 352-374.
Reder, L. M., Nhouyvanisvong, A., Schunn, C. D., Ayers, M. S., Angstadt, P., & Hiraki, K. (2000). A mechanistic account of the mirror effect for word frequency: a computational model of remember-know judgments in a continuous recognition paradigm. *J Exp Psychol Learn Mem Cogn, 26*(2), 294-320.
Reder, L., Anderson, J., Bjork, R. (1974). A semantic interpretation of encoding specificity. *Journal of Experimental Psychology 102 (4)*, 648–656.
Rosenbaum, R. S., Moscovitch, M., Foster, J. K., Schnyer, D. M., Gao, F., Kovacevic, N., . . . Levine, B. (2008). Patterns of autobiographical memory loss in medial-temporal lobe amnesic patients. *J Cogn Neurosci, 20*(8), 1490-1506.

Rotello, C. M., Macmillan, N. A., & Reeder, J. A. (2004). Sum-difference theory of remembering and knowing: a two-dimensional signal-detection model. *Psychol Rev, 111*(3), 588-616.
Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends Cogn Sci, 11*(6), 251-257.
Rugg, M. D., & Yonelinas, A. P. (2003). Human recognition memory: a cognitive neuroscience perspective. *Trends Cogn Sci, 7*(7), 313-319.
Rugg, M. D., Doyle, M. C. (1992). Event-related potentials and recognition memory for low- and high-frequency words. *Journal of Cognitive Neuroscience, 4*, 69–79.
Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A. M., Birch, C. S., & Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature, 392*, 595–598.
Rugg, M. D., Coles, M. G. H. (1995). The ERP and cognitive psychology: conceptual issues. In: Rugg, M. D., and Coles, M. G. H., (eds.) *Electrophysiology of mind: event-related brain potentials and cognition*, vol. 25. Oxford: Oxford Psychology Press, p. 27–39.
Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM - retrieving effectively from memory. *Psychon Bull Rev, 4*(2), 145-166.
Sikström, S. (2001). The variance theory of the mirror effect in recognition memory. *Psychon Bull Rev, 8*(3), 408-438.
Skinner, E. I., & Fernandes, M. A. (2007). Neural correlates of recollection and familiarity: a review of neuroimaging and patient data. *Neuropsychologia, 45*(10), 2163-2179.
Slotnick, S. D. (2010). "Remember" source memory ROCs indicate recollection is a continuous process. *Memory, 18*(1), 27-39.
Slotnick, S. D., & Dodson, C. S. (2005). Support for a continuous (single-process) model of recognition memory and source memory. *Mem Cognit, 33*(1), 151-170.
Smith, D. G., & Duncan, M. J. (2004). Testing theories of recognition memory by predicting performance across paradigms. *J Exp Psychol Learn Mem Cogn, 30*(3), 615-625.
Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: applications to dementia and amnesia. *J Exp Psychol Gen, 117*(1), 34-50.
Squire, L. R., Wixted, J. T., & Clark, R. E. (2007). Recognition memory and the medial temporal lobe: a new perspective. *Nat Rev Neurosci, 8*(11), 872-883.
Stark, S., Gordon, B., & Stark, C. (2008). A case study of amnesia: exploring a paradigm for new semantic learning and generalization. *Brain Inj, 22*(3), 283-292.
Starns, J. J., Rotello, C. M., & Ratcliff, R. (2012). Mixing strong and weak targets provides no evidence against the unequal-variance explanation of ZROC slope: a comment on Koen and Yonelinas (2010). *J Exp Psychol Learn Mem Cogn, 38*(3), 793-801.
Stenberg, G., Hellman, J., & Johansson, M. (2008). The memorability of names and the divergent effects of prior experience. *European Journal of Cognitive Psychology, 20*(2), 312-345.
Stenberg, G., Hellman, J., Johansson, M., & Rosen, I. (2009). Familiarity or conceptual priming: event-related potentials in name recognition. *J Cogn Neurosci, 21*(3), 447-460.
Toth, J. P., & Hunt, R. R. (1999). Not one versus many, but zero versus any: Structure and function in the context of the multiple memory-systems debate. In J. K. Foster and M. Jelicic (Eds.), *Memory: System, process, or function?* (pp. 232-272), Oxford University Press.
Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), Organization of memory. New York: Academic Press.
Tulving, E. (1983). Elements of episodic memory. *New York: Oxford University Press.*

Tulving, E. (1985). How many memory system are there? *American-Psychologist, 40*(4), 385-398.
Tulving, E. (2001). Episodic memory and common sense: how far apart? *Philos Trans R Soc Lond B Biol Sci, 356*(1413), 1505-1515.
Tulving, E. , & Thompson, D. M. (1973). Encoding specificity and retrieval process in episodic process. *J Exp Psychol, 87*, 353-373.
Tulving, E., Hayman, C. A., & Macdonald, C. A. (1991). Long-lasting perceptual priming and semantic learning in amnesia: a case experiment. *J Exp Psychol Learn Mem Cogn, 17*(4), 595-617.
Tulving, E., Kapur, S., Craik, F. I. M., Moscovitch, M., & Houle, S. (1994). Hemispheric encoding/retrieval asymmetry in episodic memory: Positron emission tomography findings. *Proceedings of National Academy of Science. USA, 91*, 2016-2015.
Van Petten, C., Senkfor, A. J., Newberg, W. M. (2000). Memory for drawings in locations: spatial source-memory and event-related potentials. *Psychophysiology, 37*, 551–564.
Vargha-Khadem, F., Gadian, D. G., Watkins, K. E., Connelly, A., Van Paesschen, W., & Mishkin, M. (1997). Differential effects of early hippocampal pathology on episodic and semantic memory. *Science, 277*(5324), 376-380.
Verfaellie, M., Keane, M. M., & Johnson, G. (2000). Preserved priming in auditory perceptual identification in Alzheimer's disease. *Neuropsychologia, 38*(12), 1581-1592.
Voss, J. L., Lucas, H. D., Paller, K. A. (2010). Conceptual priming and familiarity: two different expressions of memory during recognition testing with distinct neurophysiological correlates. *Journal of Cognitive Neuroscience, 22*, 2638-2651.
Voss, J. L., Lucas, H. D., Paller, K. A. (2012). More than a feeling: pervasive influences of memory without awareness of retreival. *Cognitive Neuroscience, 3-4*, 193-207.
Voss, J. L., & Paller, K. A. (2009). Remebering and knowing: electrophysciological distinctions at encoding but not retrieval. *Neuroimage, 46*, 280-289.
Wais, P. E. (2008). FMRI signals associated with memory strength in the medial temporal lobes: a meta-analysis. *Neuropsychologia, 46*(14), 3185-3196.
Wais, P. E., Mickes, L., & Wixted, J. T. (2008). Remember/know judgments probe degrees of recollection. *J Cogn Neurosci, 20*(3), 400-405.
Weingartner, H., Adefras, W., Eich, J. E., Murphy, D. L. (1976). Encoding-imagery specificity in alcohol state-dependent learning. *Journal of Experimental Psychology: Human Learning and Memory 2 (1)*, 83–87.
Westmacott, R., & Moscovitch, M. (2003). The contribution of autobiographical significance to semantic memory. *Mem Cognit, 31*(5), 761-774.
Wilding, E. L., Doyle, M. C., Rugg, M. D. (1995). Recognition memory with and without retrieval of context: an event-related potential study. *Neuropsychologia, 33*, 743–767.
Wilding, E. L., Rugg, M. D. (1996). An event-related potential study of recognition memory with and without retrieval of source. *Brain, 119*, 889–905.
Wilding, E. L. (2000). In what way does the parietal ERP old/new effect index recollection?. *International Journal of Psychophysiology*, 35(1), 81-87.
Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychol Rev, 114*(1), 152-176.
Wixted, J. T. (2009). Remember/Know judgments in cognitive neuroscience: An illustration of the underrepresented point of view. *Learn Mem, 16*(7), 406-412.
Wixted, J. T., & Mickes, L. (2010). A continuous dual-process model of remember/know judgments. *Psychol Rev, 117*(4), 1025-1054.

Wixted, J. T., & Stretch, V. (2004). In defense of the signal detection interpretation of remember/know judgments. *Psychon Bull Rev, 11*(4), 616-641.
Woodruff, C. C., Hayama, H. R., & Rugg, M. D. (2006). Electrophysiological dissociation of the neural correlates of recollection and familiarity. *Brain Research, 1100*, 125–135.
Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: evidence for a dual-process model. *J Exp Psychol Learn Mem Cogn, 20*(6), 1341-1354.
Yonelinas, A. P. (2001a). Components of episodic memory: the contribution of recollection and familiarity. *Philos Trans R Soc Lond B Biol Sci, 356*(1413), 1363-1374.
Yonelinas, A. P. (2001b). Consciousness, control, and confidence: the 3 Cs of recognition memory. *J Exp Psychol Gen, 130*(3), 361-379.
Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46*, 441–517.
Yonelinas, A. P., & Parks, C. M. (2007). Receiver operating characteristics (ROCs) in recognition memory: a review. *Psychol Bull, 133*(5), 800-832.
Yovel, G., & Paller, K. A. (2004). The neural basis of the butcher-on-the-bus phenomenon: when a face seems familiar but is not remembered. *Neuroimage, 21*, 789-800.
Yu, S. S., & Rugg, M. D. (2010). Dissociation of the electrophysiological correlates of familiarity strength and item repetition. *Brain Research, 1320*, 74–84.
Zion-Golumbic, E., Kutas, M., & Bentin, S. (2010). Neural dynamics associated with semantic and episodic memory for faces: evidence from multiple frequency bands. *J Cogn Neurosci, 22*(2), 263-277.

Paper I

EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY
2008, 20 (2), 312–345

Psychology Press
Taylor & Francis Group

The memorability of names and the divergent effects of prior experience

Georg Stenberg and Johan Hellman
Kristianstad University, Sweden

Mikael Johansson
Lund University, Sweden

Pre-experimental familiarity can have paradoxical effects on episodic memory. Knowledge of the stimulus domain usually enhances memory, but word frequency—a presumed correlate of prior experience—is negatively related to recognition accuracy. The present study examined episodic recognition of names and its relation to two measures of pre-experimental knowledge, name frequency, and fame. Frequency was operationalised as the number of hits in a national telephone directory, and fame as hits on national mass media websites. Recognition accuracy was increased by fame, but diminished by frequency. Four experiments confirmed the findings, using yes/no recognition, ROC curves, and remember-know paradigms. Hit rates were consistently more strongly influenced by fame than by frequency, whereas the reverse was true for false alarm rates. These dissociations suggest that two different forms of semantic memory, specific and nonspecific knowledge, interact with episodic memory in separate ways.

The distinction between episodic and semantic memory has become well established, but research about the two phenomena largely proceed on separate tracks. Semantic memory studies are typically concerned with the structure of conceptual organisation, whereas episodic memory research cares mainly about the processes of encoding and retrieval. Yet there is considerable interdependence between the two branches, because events encoded in episodic memory are normally interpreted against a background of semantic knowledge, and when such foreknowledge is lacking, remembering suffers (Bartlett, 1932).

Correspondence should be addressed to Georg Stenberg, Dept. of Psychology, School of Behavioural Sciences, Kristianstad University, SE-291 88 Kristianstad, Sweden. E-mail: georg.stenberg@bet.hkr.se

http://www.psypress.com/ecp DOI: 10.1080/09541440701398724

As a first approximation to a generalisation, frequent prior experience of to-be-studied items should enhance episodic memory. From the perspective that memory is adapted to the environment, it is clearly rational to preserve in memory those items most likely to be encountered again (Anderson & Schooler, 1991, 2000; Dennis, 1995). Thus, items about which semantic knowledge has been amassed are more prone to be remembered after episodic encounters than less known items. In an early study (Allen & Garton, 1968), physics students recognised physics words from a list better than did arts students, but both groups of students recognised physics words better than common words. Similarly, in a more recent study (Chalmers, Humphreys, & Dennis, 1997), computer science students showed better episodic recognition of rare computer science terms than psychology students, although both groups showed an advantage for low-frequency over high-frequency words. The finding that relatively rare words are better recognised than relatively common words is a consistent and widely reproduced finding (reviewed by Chalmers & Humphreys, 1998; Reder et al., 2000), but it flies in the face of the generalisation that frequent prior experience enhances memory.

The *word frequency effect* has been a challenge for memory theories in more ways than one. Because most early models did not take pre-experimental experience into account, there was no explanation of the frequency effect. Second, if prior familiarity *was* explicitly addressed, theories predicted that it would raise *both* the hit rate *and* the false alarm rate. Instead, the typical finding is that hit rates are raised and false alarm rates are lowered. This pattern is often called the *mirror effect*, because the old and the new distributions are thought of as moving away from the criterion in opposite directions (Glanzer & Adams, 1985, 1990; Glanzer, Adams, Iverson, & Kim, 1993).

Word frequency, as measured by the probability of a word's occurrence in (usually newspaper) text, is expected to reflect experiential frequency, i.e., probability of encountering the word in daily life. That assumption has been called into question (Estes & Maddox, 2002). A more closely controlled manipulation of pre-experimental experience can be accomplished in a three-phase experiment, where stimuli are presented in a familiarisation phase before the usual study–test procedure. With both nonverbal and verbal materials, repeated presentations preceding the study and test phases have been seen to increase both hit rates and false alarm rates, sometimes leading to a net decline in sensitivity, d' (Estes & Maddox, 2002; Maddox & Estes, 1997).

The effect brought about by word frequency is different from prior familiarisation, although the specific functional relationship can vary depending on the range of frequencies included. If the range is wide and includes very low frequencies, the typical finding is an inverted-U relationship, with maximum sensitivity at moderately low frequencies. The peak is reached by the combined effect of both high hit rates and low false alarm rates. In view of these divergent trends, Estes and Maddox (2002) concluded that the word frequency effect is a misnomer, because it reflects a quality different from experiential frequency, probably *lexicality* in a wide sense, which tends to be confounded with frequency.

The model used by Estes and Maddox (2002) belongs to the class of global matching models for recognition memory (Clark & Gronlund, 1996). A competing type of models, the *dual-process* class, has made a different interpretation. Two processes are available in recognition, according to these theories: recollection and familiarity, and they contribute in different ways to the overall word frequency effect (Arndt & Reder, 2002; Reder et al., 2000). The hit rate part reflects the fact that low-frequency words are more often recollected. The false alarm part can be ascribed to the greater familiarity of high-frequency words. Because there are two different contributions, they can be pried apart by experimental manipulations. Thus, if task demands necessitate extraordinarily fine discriminations to be made between targets and similar distractors, the data will show responses consistent with a high degree of recollection, reflected in the shape of the ROC curves (Arndt & Reder, 2002). If remember–know responses are recorded, high-frequency words will attract more know-responses than low-frequency words do, in keeping with their higher familiarity (Reder et al., 2000).

The present study is concerned with the kind of prior experience subsumed under the heading of general knowledge or semantic memory, and the effects this experience has on episodic memory. We use proper names as the stimulus material, these being the object of everyday semantic knowledge (as well as an oft-lamented source of memory lapses). We propose that knowledge of names takes two forms, one distinctive, such as knowing that Björn Borg is the name of a celebrated tennis player, and one nondistinctive, such as knowing that Tom Jones is a common name and Engelbert Humperdinck is not. These two forms of semantic memory bear a structural resemblance to two forms of episodic memory, recollection and familiarity. Furthermore, they interact with episodic memory in radically different ways. Distinctive semantic knowledge supports episodic memory, whereas the nondistinctive form may interfere with it.

Distinctive semantic memory shares with recollection a relative richness of associated detail; as applied to proper names it singles out a particular bearer of the name, and brings to mind known facts about that person. The name of a celebrity is associated with that person's looks, achievements, public appearances, etc., all of which can help to encode an episodic encounter with the name, such as seeing it in the newspaper. Nondistinctive memory, on the other hand, brings to mind a sense of familiarity, a sense of many previous encounters with the name, without any particular one coming to the fore. Frequent names—such as Smith and Jones—often evoke such a sense of familiarity without bringing any particular bearer to mind. We aim to show that such familiarity is detrimental to episodic recognition memory. In that respect, the *name frequency effect*, which we aim to document, is a close relative of the word frequency effect. As mentioned above, it has been proposed that the word frequency effect is not really a result of frequency of encounters with the words in daily life (Estes & Maddox, 2002). Instead, lexicality, with which frequency may be confounded, could be the causative factor. Nonetheless, the present study provides further indications that the frequency effect is genuine, by demonstrating a parallel effect, using a different stimulus material.

Names (i.e., first name plus last name) that were used in this study varied orthogonally along two dimensions: frequency and celebrity. Both were meant to reflect environmental quantities. With *frequency* we refer to the relative number of persons bearing the name, by *celebrity* we mean the probability of the name being mentioned in the media. Operationally, frequency was measured as the number of hits in a computerised search of the national telephone directory. Celebrity was similarly defined as the number of hits in a search of the Internet pages of national news media. Conceptually as well as empirically, these are independent criteria, and four groups could be formed by cross-classifying high and low groups. Examples are given later, and the full lists of names are available on the Internet.[1] In the experiments, names were presented for study visually, and tested for recognition shortly thereafter.

We aimed to show that celebrity and frequency had opposite-sign effects on recognition accuracy, but we also wanted to show that they affected at least partly different memory systems. Our expectation was that distinctive semantic knowledge (correlated with celebrity) would feed into the episodic *recollection* system. By providing material for detailed encoding, it would engender context-rich memories.

Nondistinctive semantic knowledge (correlated with frequency), on the other hand, would feed into the episodic *familiarity* system, causing confusion in the process. Because familiarity is context-free by definition, it needs to be attributed to a source, and if several sources are possible, confusion may arise. The familiarity arising from frequent occurrences in the pre-experimental environment cannot be easily distinguished from the familiarity arising from study within the experiment. Because of this, frequency raises the noise level in the recognition process, and makes the signal difficult to distinguish.

[1] http://www.stenberg.ys.se/Projects/Names/Names.htm

Chalmers and Humphreys (1998) distinguished between generalised and episode specific strength in episodic recognition of high- and low-frequency words. Both factors were varied experimentally, generalised strength by frequency of presentation, and episode specific strength by the presentation of definitions in the familiarisation phase. Whereas making words distinctive by supplying definitions supported recognition accuracy, mere familiarisation was of doubtful value for recognition accuracy; in some conditions the effect was detrimental. The distinction we wish to make is similar to that of Chalmers and Humphreys, although distinctive semantic memory in our sense need not be episode specific. Knowledge concerning, e.g., Björn Borg is genuinely semantic in the sense of being accumulated over many instances. Events in which information about him was presented have themselves faded away from memory, leaving only the information behind, much as ancestors, long dead, have contributed the genetic code of the living.

Undoubtedly, episodes may be remembered still by many who witnessed Borg's Wimbledon victories, and they may even form personally relevant memories, perhaps charged with emotion and autobiographical significance. The point we wish to make is that episodes are not a *necessary* component of distinctive semantic memory. Knowledge of famous people can be detailed and individuating without being woven into the fabric of our own lives.

The distinction between autobiographically relevant and irrelevant knowledge of famous people has been elucidated by Moscovitch and colleagues (Westmacott & Moscovitch, 2003). They showed that celebrities that were associated with events in participants' own lives were better remembered in an episodic recognition experiment than celebrities not so associated. Furthermore, self-relevance also improved performance in tests of semantic memory. Westmacott & Moscovitch's results demonstrate the interdependence of semantic and episodic memory. We wish to make a related although different proposal, namely that a dividing line runs within semantic memory itself, irrespective of autobiographical association. Both fame and frequency are likely to increase the probability that a name evokes a personal memory, yet they have—as we purport to show—radically different effects on within-experiment episodic memory.

EXPERIMENT 1

The first experiment assessed the effects of distinctive and general semantic memory on episodic memory, using a pool of names with which participants had varying and measurable foreknowledge.

Method

Participants

Forty-seven students (34 women) participated, and were compensated with a lunch voucher. Ages ranged from 16 to 40, with an average of 24 years. They were randomly allotted to two orienting tasks, resulting in 27 participants in the celebrity orienting task, and 20 in the frequency orienting task.

Procedure

Participants were tested in groups of 2–16 at a time. The experiment took place in a laboratory, where participants were seated in separate booths, each with a computer, on which stimuli were presented using E-prime software. After a brief oral instruction and some on-screen instructions, participants ran the experiment at their own pace. The whole session took about half an hour.

The present experiment was divided into three study–test cycles, each presenting 32 names for study. Studied names reappeared in the test mixed with an equal number of distractors. Assignment of names to the study set or the distractor set was randomly and independently determined for each participant, as was the presentation order. The test phase followed the study phase without delay.

During study, each stimulus was preceded by a 1 s fixation cross. The name was displayed, centred on the screen, for 2 s, and during this time window a response was to be given to the orienting question (presented at the top of the screen as a reminder): either "Is this person famous?" or "Is this name frequent?" (The latter question had been specified in the instruction text to mean "Are there more than 10 bearers of that name in Sweden?").

During the test phase, names were presented until a response was given or until 4 s had elapsed, whichever happened first. After the response, a brief (0.5 s) feedback concerning correctness of the response was given.

Materials

A priori ratings. A set of 192 Swedish names was constructed. Names were either selected from the set of those currently (late 2005) popular in the media, or combined (first name plus last name) using frequency tables provided by the national census bureau, Statistics Sweden. The experimenters, when constructing the stimulus material, judged each name as either Famous or Nonfamous, and either Frequent or Infrequent. Forty-eight names of each type were selected. Examples of common, famous names were: Göran Persson, Ingmar Bergman, and Björn Borg; and of uncommon, famous names: Ingvar Kamprad, Greta Garbo, and Zlatan Ibrahimovic. Common, nonfamous names were, e.g., Maria Axelsson, Sven Holmgren, and Gustav Eklund, and rare, nonfamous names, Ernfrid Hammar, Hildegard Sten, and Guje Gagner. Celebrity names included ones whose claim to fame extended over decades (Garbo, Bergman), as well as others of more recent renown. The present set, and an expanded, second set are available at: http://www.stenberg.ys.se/Projects/Names/Names.htm.

Internet searches. To verify the judgements, names were checked for frequency by looking up each name in the Swedish, nationwide telephone directory (www.eniro.se), and noting the number of hits. This number was log transformed, and used as the variable Frequency, which was dichotomised into Frequent and Infrequent.

Similarly, *celebrity* was checked by making site-specific lookups via the Google search engine. Each name was searched at six Swedish websites, affiliated with important media: four national newspapers (www.dn.se, www.svd.se, www.expressen.se, www.aftonbladet.se) and two television networks (www.svt.se and www.tv4.se). Searches were made by a Visual Basic program, using a programming interface published by Google (www.google.com). The number of hits was added across sites, and log transformed. Finally, the variable was dichotomised into Famous and Nonfamous.

Participant ratings. In addition to the search data, participants were invited to rate the names for frequency or celebrity. Each rating task was allotted to one half of the participants as an orienting task in the study phase. Apart from providing validation of the stimulus classification, it also served to examine whether attention directed towards one dimension of the stimuli would affect memory for the names.

A priori ratings of celebrity and frequency were confirmed by the ratings given by participants and the pattern of hits in Internet searches. Agreement (Kendall's tau) was .86 with participant classification, and .98 with Internet hits, and between the latter two the correlation was .84 (see Table 1).

Internet hits in the phone directory was unrelated to hits on the media sites ($r = .04$), but both were correlated with the total sum of Google hits, a possible indicator of experiential frequency ($r = .43$ for the phone directory, and $r = .71$ for the media; $n = 192$).

Results

To allow generalisation across both subjects and items, two sets of analyses were performed (Clark, 1973), one with subjects, and the other with items, as

TABLE 1
Agreement between classification of stimuli by a priori judgements (columns) and empirical classification by Internet search (rows, upper half), and participant judgements (rows, lower half)

		Frequent		*Infrequent*		
		Famous	*Nonfamous*	*Famous*	*Nonfamous*	*Total*
Internet search						
Frequent	Famous	47	3			50
	Nonfamous	1	45			46
Infrequent	Famous			47		47
	Nonfamous			1	48	49
Total		48	48	48	48	192
Participant judgements						
Frequent	Famous	43	3	4		50
	Nonfamous	1	43	1	1	46
Infrequent	Famous	4	1	41		46
	Nonfamous		1	2	47	50
Total		48	48	48	48	192

source of the random error term. Although this analysis strategy has been questioned as a general practice (Raaijmakers, Schrijnemakers, & Gremmen, 1999), it is arguably necessary with the present data set, because random selection of items was possible only within conditions (Frequency and Celebrity), not across condition boundaries. Furthermore, analysis by items permitted examination of issues not otherwise accessible. In the subjects-based analysis, the a priori classification of the names was used, and in the item-based analysis, we used the empirical classification derived from Internet searches. Thus, we could perform regression analyses, predicting item hit rates and false alarm rates from our ratio-scaled Internet data.

By subjects. Hit rates and false alarm rates were recorded for each of the four *a priori* stimulus classes, averaged over items for each participant, and d' was computed (Snodgrass & Corwin, 1988). It was subjected to a $2 \times 2 \times 2$ analysis (Frequency $\times$ Celebrity $\times$ Orienting Task), the first two being within-participant factors, and the third a between-participant factor. Orienting Task had no effect, alone or in interaction, and it will not be further mentioned.

The d' measure (see Figure 1) showed a main effect of Frequency, because uncommon names were better recognised, $F(1, 45) = 42.19, p < .001, \eta_p^2 = .48$. Famous names were much better remembered than nonfamous, $F(1, 45) = 238.96$, $p < .001$, $\eta_p^2 = .84$. There was also an interaction between Celebrity and Frequency, $F(1, 45) = 6.24$, $p = .016$, $\eta_p^2 = .12$, due to a potentiated frequency effect among the celebrities.

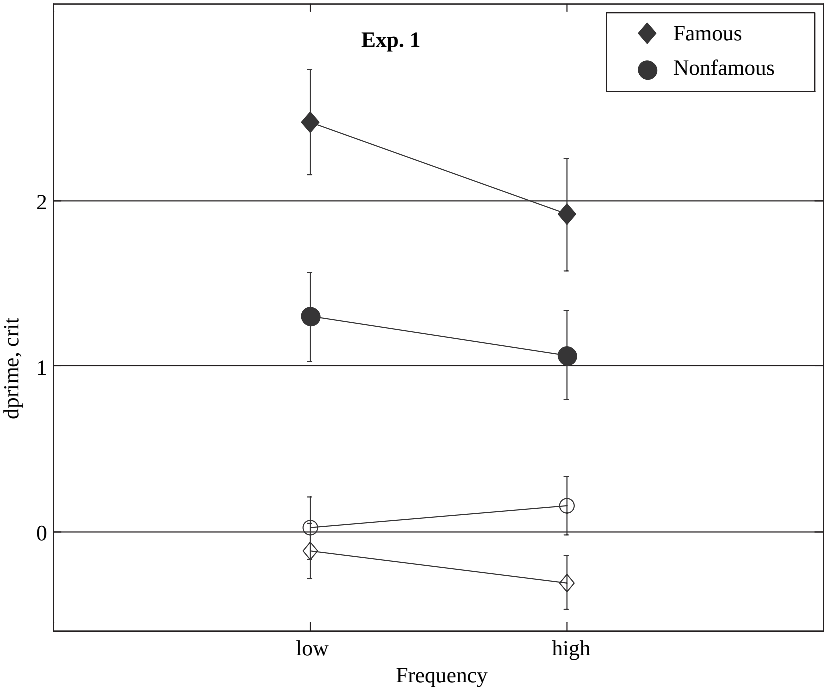


Figure 1. Experiment 1: Values of d' (filled symbols) and C (unfilled symbols); circles: famous names; diamonds: nonfamous. Frequency is on the x-axis. Error bars show ± 1 standard error.

The criterion, C, was lower for famous names than for nonfamous, resulting in a Celebrity effect, $F(1, 45) = 67.38$, $p < .001$, $\eta^2_p = .59$, and an interaction, Celebrity × Frequency, $F(1, 45) = 27.35$, $p < .001$, $\eta^2_p = .37$ (Figure 1).

Hit rates showed an effect of Frequency, $F(1, 46) = 11.79$, $p = .001$, $\eta^2_p = .20$, and a large effect of Celebrity, $F(1, 46) = 231.20$, $p < .001$, $\eta^2_p = .83$, and an interaction, $F(1, 46) = 11.50$, $p = .001$, $\eta^2_p = .20$.

False alarm rates were affected by Frequency, $F(1, 46) = 24.41$, $p < .001$, $\eta^2_p = .35$, and by Celebrity, $F(1, 46) = 14.19$, $p < .001$, $\eta^2_p = .24$, as well as by an interaction, $F(1, 46) = 25.83$, $p = .001$, $\eta^2_p = .36$. Averages are given in Table 2.

Effect sizes for the effects of Frequency and Celebrity are also given in Table 2. There has been discussion as to which index of effect size is preferable in within-subjects designs, the η^2_p or the η^2_G, the latter being more comparable to between-subjects designs (Bakeman, 2005). We present both, to allow comparisons; the former in text and the latter in tables.

TABLE 2
Hit rates and false alarm rates in Experiment 1

	Means					
	Low frequency		*High frequency*		*Effect sizes, η^2_G*	
	Nonfamous	*Famous*	*Nonfamous*	*Famous*	*Frequency*	*Celebrity*
HR	.73	.90	.65	.90	.03	.63
FAR	.25	.14	.25	.26	.11	.08
HR (by items)	.74	.90	.65	.88	.07	.52
FAR (by items)	.25	.14	.24	.27	.04	.02
Mirror effect patterns	HR1	HR2	FAR2	FAR1		
Celebrity	.90	.69	.25	.20		
Frequency	.82	.78	.26	.20		

The lower part of the table verifies the mirror effect pattern by showing hit rates and false alarm rates collapsed across high and low levels of Frequency and Celebrity, respectively. The means are enumerated from left to right in the expected order of magnitude, under the assumption of a mirror effect. As the table shows, the data conformed to the expected pattern, inasmuch as there were mirror effects for both Frequency and Celebrity.

By items. Hits and false alarms were recorded for each stimulus, averaged over participants for each item, and item-wise hit rates and false alarm rates were computed. They were subjected to a 2×2 analysis (Frequency $\times$ Celebrity), both being between-items factors.

There was a reliable effect of Celebrity, $F(1, 188) = 206.04$, $p < .001$, $\eta^2_p = .52$, and of Frequency, $F(1, 188) = 13.92$, $p < .001$, $\eta^2_p = .07$ on hit rates. The interaction was also significant, $F(1, 188) = 6.53$, $p = .011$, $\eta^2_p = .03$.

False alarm rates showed an effect of Frequency, $F(1, 188) = 8.70$, $p = .004$, $\eta^2_p = .04$, that was stronger than the effect of Celebrity, $F(1, 188) = 4.35$, $p = .038$, $\eta^2_p = .02$. There was also an interaction, $F(1, 188) = 11.42$, $p = .001$, $\eta^2_p = .06$, due to particularly low false alarm rates for infrequent, famous names, compared to the other three groups.

Using the full range of predictor variables, instead of dichotomies, a multiple regression analysis was performed. First, *hit rate* was used as the dependent variable. Four potential predictors were entered into a stepwise regression: (a) the endorsement rate in the frequency orienting task (*freq_calls*), (b) the endorsement rate in the celebrity orienting task (*celeb_calls*), (c) the number of hits in the Internet search of the telephone directory (*freq_hits*), and (d) the number of hits in the Internet search of the media sites (*celeb_hits*). The latter two were log transformed for normality. The stepwise regression procedure settled for a model with three predictors:

Celeb_calls: $\beta = .52$, $t(188) = 8.40$, $p < .001$; *freq_calls*: $\beta = -.23$, $t(188) = -4.88$, $p < .001$; *celeb_hits*: $\beta = .28$, $t(188) = 4.58$, $p < .001$.

The same analysis was applied to the *false alarm rate*. Thus, a stepwise regression procedure was offered the same four potential predictors. Only one was selected as significant: *freq_hits*: $\beta = .18$, $t(188) = 2.51$, $p = .013$.

Thus, whereas hit rates were predicted by indicators of both frequency and celebrity (and more so by celebrity), false alarm rates were predicted by frequency alone.

Discussion

In keeping with the premise that prior experience furthers memory, famous names were retained much better than nonfamous names. However, completing the paradox of pernicious foreknowledge, frequent names were not well remembered at all; indeed, they fared far worse than very unusual names. The frequency effect was boosted in the group of famous names, but it was significant for famous and nonfamous alike.

Some of the infrequent names were combined from relatively rare constituent parts, and the resulting first + last name pairs could have had an unusual look and sound. To ascertain whether the frequency effect could be ascribed to this bizarreness aspect, a modified stimulus material was used in later experiments. It can be noted already at this point, however, that the frequency effect was in fact stronger among the celebrities, this being the basis of a significant interaction effect. In the famous names, any potential bizarreness would have been eroded by constant wear and use in the media, and to most native speakers these names would appear to be familiar household names.

Although both factors had quite marked effects on d', separate influences could be noted on hit rates and false alarm rates. The impact of Celebrity on hit rates was huge, and that of Frequency paled by comparison. False alarms rates, on the other hand, were affected by Frequency, more so than by Celebrity, although both influences were relatively weak. It can be noted that the distinct patterns of effects on HR and FAR are some of the most important characteristics used to distinguish between familiarity and recollection (Reder et al., 2000).

EXPERIMENT 2

The purpose driving our further research in this area was, apart from replicating the basic findings, a wish to clarify the relation between the two dimensions of semantic memory and similar dimensions of episodic memory. In particular, we wished to make contact with the flourishing research on familiarity and recollection, which has proposed methods of separating and measuring the two. An often used method is the *Remember–Know* paradigm, in which participants are questioned about the introspective quality of their memory. In the present context, high-celebrity names can be expected to elicit *Remember* responses. Analogously, frequency could possibly affect the rate of *Know* responses, perhaps especially those given erroneously, i.e., Know false alarms. We therefore adapted standard Remember/Know instructions (Rajaram, 1996) to the name memory task, which was otherwise presented as in Experiment 1. We anticipated some difficulty for the participants in performing the remember/know task with this particular material, arising from the possible confusion of pre-experimental familiarity with familiarity engendered within the experiment. In the instructions, we therefore emphasised that the subjective quality to which the label "Know" applied had nothing to do with previous knowledge (acquired, e.g., through the media), and that we asked participants to separate this way of "knowing" from the kind we wanted them to report, i.e., the familiarity produced by an earlier encounter within the experiment.

Method

Participants

Thirty-five students (26 women) participated in exchange for a lunch voucher. Ages ranged from 19 to 47, with an average of 23 years.

Procedure

Participants were tested in groups of 2–16 at a time. The experiment took place in a laboratory, where participants were seated in separate booths, each with a computer. After a brief oral instruction, they were given further on-screen instructions, and then ran the experiment at their own pace. This experiment was interleaved with a different, unrelated experiment, such that study of all (64) to-be-remembered names came first, followed by the other experiment (an Iowa gambling task), and finally a memory test for all the names (128). The retention period—about 10 min—was thus filled with a distracting task, but no names or other verbal material appeared in it. The whole session took about 40 min.

During study, each name was shown for 2 s, and no overt task was assigned during this period, except to memorise. In the test block, each name was presented along with two on-screen buttons, marked "Yes" and "No", to be mouse-clicked during a 5 s period in response to the question "Did you see this name before in the experiment?". In case of a "Yes" response, a three-button selection screen followed, with choices marked "Jag minns det" ("I remember"), "Det känns bekant" ("I know"), and "Jag gissar"

("I guess"). These alternatives had been extensively explained at the outset, using a modified version of the Rajaram (1996) instructions. "No" responses were not followed by any further selection.

Materials

An amended and expanded stimulus set of 288 names was used in this experiment. The main changes from the previous set were (a) the inclusion of 96 new names; (b) substitution of some infrequent, nonfamous names that could give rise to a bizarreness effect—all names (first + last name combination) had to have at least one directory-listed bearer as a requisite for inclusion (exceptions were made for nonlisted or deceased famous persons such as Greta Garbo); (c) an update of the fame and frequency data was performed through a renewed web search (February 2006), about 6 months later than the previous one—for the set of identical items in the two sets ($n = 132$), correlations were .95 and .87 for the phone directory hits and the media hits, respectively.

The number of phone directory hits was unrelated to the number of media hits ($r = .01$), but both were positively correlated with a third variable, which can be thought as a proxy for experiential frequency, the total number of Google hits ($r = .32$ and $r = .82$, respectively, all variables log transformed).

The selection of 64 targets and 64 distractors out of the 288 item pool was made randomly and independently for each participant, with the constraint that the four types of names be equally represented.

Results

As in the previous experiment, analysis was performed both by-subjects and by-items. Starting with the former, we made conventional analyses of d', hit rates and false alarm rates, all in a 2×2 design (Frequency $\times$ Celebrity).

The remember–know data yielded two sets of variables, the remember rate (r), computed as the proportion of remember responses out of all old items, and the IRK-know rate (Table 3). The latter was computed as suggested by Jacoby and colleagues (Jacoby, Jones, & Dolan, 1998; Kelley & Jacoby, 1998), i.e., as the proportion of know responses out of targets not given a "remember" response. These variables, r and IRK_k, were subjected to a 2×2 analysis (Frequency $\times$ Celebrity). Additionally, we computed rFA and kFA, i.e., the proportion of false alarms given remember and know responses. The rate of false remember responses has been the focus of theoretical interest (Wixted & Stretch, 2004), and the rate of false know responses interested us because we suspected that frequency might have an influence on it.

TABLE 3
Hit rates and false alarm rates in Experiment 2

	Means					
	Low frequency		*High frequency*		*Effect sizes, η^2_G*	
	Nonfamous	*Famous*	*Nonfamous*	*Famous*	*Frequency*	*Celebrity*
HR	0.66	0.89	0.63	0.82	0.05	0.47
HR (by items)	0.68	0.89	0.62	0.83	0.02	0.27
r	0.46	0.79	0.32	0.68	0.20	0.68
IRK_k	0.54	0.86	0.41	0.76	0.15	0.60
FAR	0.16	0.11	0.22	0.22	0.17	0.02
FAR (by items)	0.14	0.12	0.23	0.22	0.07	0.00
rFAR	0.04	0.02	0.04	0.06	0.05	0.01
kFAR	0.07	0.05	0.10	0.11	0.10	0.00
Mirror effect patterns	HR1	HR2	FAR2	FAR1		
Celebrity	.86	.65	.19	.17		
Frequency	.78	.73	.22	.14		

r: remember; k: know; FAR: false alarm rate; HR: hit rate; IRK: independent remember-know. The lower part of the table verifies the mirror effect pattern by showing hit rates and false alarm rates collapsed across high and low levels of Frequency and Celebrity, respectively. See note to Table 1.

By subjects. The d' measure (Figure 2) was affected by both Frequency, $F(1, 34) = 25.45$, $p < .001$, $\eta^2_p = .43$, and Celebrity, $F(1, 34) = 86.55$, $p < .001$, $\eta^2_p = .72$. There was also an interaction, $F(1, 34) = 7.55$, $p = .010$, $\eta^2_p = .18$, due to larger differences between frequent and infrequent names among the famous. Despite differences in the stimulus material, all aspects of these results were well reproduced from the previous experiment.

Participants set the criterion, C, higher for nonfamous names, again replicating the Celebrity effect from the previous experiment, $F(1, 45) = 28.87$, $p < .001$, $\eta^2_p = .46$. There was no Frequency effect and no interaction.

Hit rates showed a large effect of Celebrity, $F(1, 34) = 89.36$, $p < .001$, $\eta^2_p = .72$, only a marginal effect of Frequency, $F(1, 34) = 3.74$, $p = .062$, $\eta^2_p = .10$, and no interaction, $F < 1$.

False alarm rates were affected by Frequency, $F(1, 34) = 20.66$, $p < .001$, $\eta^2_p = .38$, with no Celebrity effect and no interaction (both $p > .10$). More false alarms were made to frequent than to infrequent names.

Remember responses. Remember responses to old items were affected by both Frequency, $F(1, 34) = 21.67$, $p < .001$, $\eta^2_p = .39$, and Celebrity, $F(1, 34) = 199.59$, $p < .001$, $\eta^2_p = .85$, with no interaction.

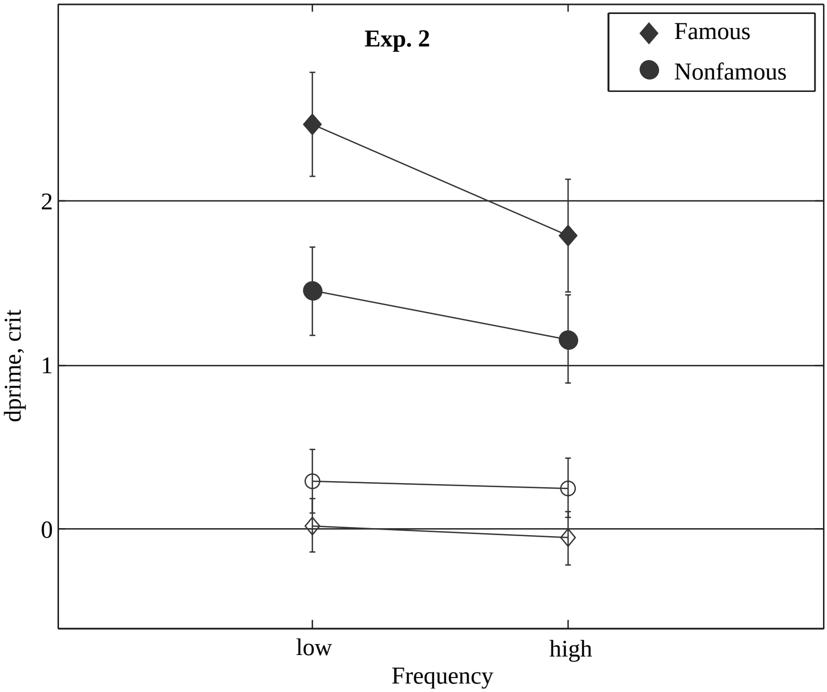


Figure 2. Experiment 2. Values of d' (filled symbols) and C (unfilled symbols); circles: famous names; diamonds: nonfamous. Frequency is on the x-axis. Error bars show ± 1 standard error.

Remember responses to new items were quite few, but interesting, considering their incompatibility with threshold theories of remembering. They were more numerous to frequent names than to infrequent, Frequency, $F(1,\ 34) = 5.45$, $p = .026$, $\eta_p^2 = .14$. Celebrity had no effect, alone or in interaction.

Know responses. Know responses to old items, computed according to IRK assumptions (Jacoby et al., 1998) reflected the same pattern as remember responses, i.e., they were affected by both Frequency, $F(1,\ 34) = 12.82$, $p = .001$, $\eta_p^2 = .27$, and Celebrity, $F(1, 34) = 141.36$, $p < .001$, $\eta_p^2 = .81$, with no interaction.

Know responses to new items reflected the pattern of all types of false alarms in being more common to frequent names than to infrequent: Frequency, $F(1,\ 34) = 7.95$, $p = .008$, $\eta_p^2 = .19$. There were no other effects.

By items. In the interest of brevity, we report only regression analyses over items, not the ANOVA resulting from dichotomising the independent variables, although the latter was also performed and showed the same pattern as the by-subjects analysis, i.e., hit rates were affected by both Celebrity and Frequency, but false alarm rates were affected by Frequency alone.

In the first regression analysis, hit rate was designated the dependent variable, and the potential predictors—(a) the log-transformed number of hits in the Internet search of the telephone directory (*freq_hits*), and (b) the log-transformed number of hits in the Internet search of the media sites (*celeb_hits*)—were introduced in a stepwise analysis. Both predictors were accepted as significant, c*eleb_hits*: $\beta = .53$, $t(285) = 10.79$, $p < .001$; *freq_hits*: $\beta = -.18$, $t(285) = -3.64$, $p < .001$.

In a similar analysis, using false alarm rate as the dependent variable, only frequency was found to be a significant predictor, *freq_hits*: $\beta = .31$, $t(285) = 5.48$, $p < .001$.

Discussion

Using a partly different name set, we replicated the basic findings from the first experiment; names are better retained if they are famous beforehand, but famous or not, common names fare worse than unusual ones. Attempting to resolve this paradox of foreknowledge, we found that hit rates and false alarm rates showed different patterns of effects. Dissociation of hits and false alarms is a hallmark of dual-process theories of recognition. Whatever the particular brand of dual-process, a shared assumption is that recollection is active only in recognising old items and can have very little effect on false alarms. The FA rate is instead shaped by familiarity, i.e., familiar items attract more false recognition responses. The quality of familiarity may or may not be helpful in making true recognition responses as well, and the particular blend of recollection and familiarity that goes into shaping the hit rate is specific to each task and context.

So far, we have traced a parallelism. Fame, like recollection, exerts its effect mainly on hit rates. Frequency, although a less potent force overall, is the dominant influence, and sometimes the only influence, on false alarm rates.

We pursued this parallelism further with remember-know methodology, speculating that the different qualities of recognition associated with frequency and celebrity might be introspectively accessible. We had some misgivings on that score, acknowledging the burden we placed upon the participants in distinguishing pre-experimental from experimental familiarity. Our misgivings proved to be well-founded. The pattern of effects was essentially the same over d', remember and know responses, suggesting that participants used the different responses mainly to express different degrees of confidence, not different qualities of experience. Indeed, one active line of criticism against remember-know research is that this is exactly what happens: "Know" expresses just a lower degree of confidence than "remember", and apart from that, qualities of experience do not enter into the decision (Hirshman & Henzler, 1998). We do not wish to enter that debate, except to affirm that the remember–know distinction was less useful for our purposes.

Having had response confidence brought to our attention, we now turn to a class of methods based on explicit confidence responses, the ROC (receiver operating characteristics) approach.

EXPERIMENT 3

Experiment 3 was performed to examine the memory characteristics of the name stimuli over a set of criteria. By inviting responses graded from high to low confidence, information can be gleaned about the shape of the ROC curves.

The ROC is a function relating hit rates to false alarm rates over a range of criteria. It has received increasing interest as an indicator of memory processes in recent years, because it gives more detailed information about the variation of accuracy when different criteria are adopted. The usefulness of ROC data for memory studies has been pointed out by many (Heathcote, 2003), although there is no complete consensus on how analysis of such data should proceed. Many approaches refer to the z-transformed ROCS (i.e., data converted by the inverse of the cumulative normal distribution), because these will be roughly linear (see Figure 3). We will also refer to these for convenience, although our analyses do not use linear regression fitting, but maximum likelihood fitting of the original data (Harvey, 2005).

Three types of parameters can be extracted: first, the distance between the old and the new distributions, which is a measure of accuracy, akin to the conventional d'. In the z-transformed graphs, this is reflected in the intercept of the regression line. There is a choice between scaling the distance by the standard deviation of the new or the old distribution or a compromise (d_a), but either way, the interpretation is relatively straightforward. With our present data, we expect the intercept to reflect the effects of frequency and celebrity on accuracy, as documented in the first two experiments.

The second parameter is the slope of the regression line. It reflects the relation between the standard deviations of the new and the old distributions, and because the new distribution is often assumed to have $\sigma = 1$, the slope is simply the inverse of the SD of the old distribution, $1/\sigma_{old}$. In some studies, slope has been seen to decrease with increased accuracy (Glanzer, Kim, Hilford, & Adams, 1999), and in others it has remained constant across conditions (Ratcliff, McKoon, & Tindall, 1994). There is still an ongoing debate about the proposed constancy-of-slopes generalisation

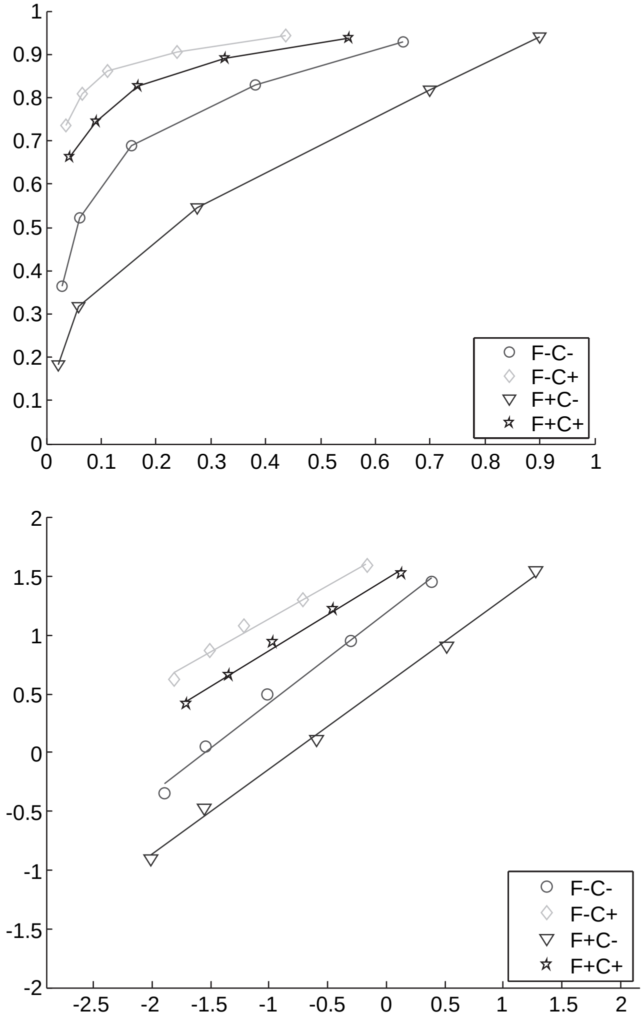


Figure 3. ROC curves (upper half) and *z*-transformed ROC-curves (lower half) for Experiment 3. Markers: circles: low frequency, low celebrity; diamonds: low frequency, high celebrity; triangles: high frequency, low celebrity; stars: high frequency, high celebrity.

(Ratcliff et al., 1994). In the present study, slopes will also be examined for effects paralleling those on intercepts, such that wherever accuracy increases, slope decreases. If, on the other hand, the constancy of slopes assumption is correct, we will see no change, or a greatly attenuated decrease. Our interest in the present study is not primarily in this matter, but in the possible dissociation of two types of prior experience. Therefore, we will attend to the effects of frequency and celebrity on slopes, and in particular whether they are parallel or divergent.

The third type of parameter is the shape of the ROC. Although difficult to capture in a single quantity, the shape reflects possible deviations from the assumed signal detection model. Some theories of recognition memory explicitly posit such deviations. Dual process theories (Yonelinas, 2002) assume that a recollection process sometimes introduces an all-or-none element into the otherwise graded and probabilistic recognition process. This produces a concave-upward bent on the otherwise linear *z*-ROC. On the other hand, other processes of an artefactual character, such as high-confidence guessing, can produce a concave-downward deviation from linearity. In fact, this type of downward curvature is arguably more common (Glanzer et al., 1999; Heathcote, 2003) than the upward curvature that signals dual-process (but see, e.g., Arndt & Reder, 2002). We will not pursue the matter further, except to note that we observed a slight downward curvature in some of our conditions (see Figures 3 and 5).

Regardless of what type of parameter one chooses to focus on, model fitting can proceed along different paths. We will sketch two modes of operation. The first—and standard—mode operates on individual data for each experimental condition separately. It produces parameter estimates that are readily accessible and can easily be submitted to conventional statistical analysis. It has the drawback that single-subject data are relatively sparse, and several cells in the data sheet may be empty. Such analysis is sensitive to noisy data and may even introduce a systematic bias (Schooler & Shiffrin, 2005).

An alternative is to operate on aggregated data from all participants, i.e., a group ROC. This eliminates problematic discontinuities and safeguards against undue influence of minor artefacts. The question arises, however, as to how the variability, or conversely, the reliability of an effect is to be estimated. A solution can be found in recently developed, computer intensive bootstrapping methods (Martinez & Martinez, 2002; Schooler & Shiffrin, 2005). By resampling with replacement from the original data, new samples are constructed in which variability reflects that of the raw data. With enough resampling (typically > 1000 times), confidence intervals for any parameter of interest can be arrived at. Thus, contrasts corresponding to ANOVA effects can be computed, and the corresponding confidence intervals can be examined as to whether they include zero or not. Alternatively, another recent statistical development can be put to use. Killeen (2005) has proposed an alternative to null hypothesis testing in p_{rep}, the probability of replication. It is defined as the probability of a new experiment of similar power arriving at an effect of the same sign. Effects conventionally described as significant typically have a p_{rep} of .90 or more.

The p_{rep} is easily computed from bootstrapping samples by tallying the number of same-sign effects.

Orthogonally to the group versus single-subject question, a choice has to be made whether to fit each condition separately (e.g., deep vs. shallow study) or all conditions simultaneously. Not all software allows the second choice, but L. Harvey's Rscore program (Harvey, 2005) permits a large number of signal conditions. The first choice is the path most often taken, but the second one uses the data more efficiently. The empirical fact speaking in favour of simultaneous fitting is the finding that people tend to use the same criteria in judging stimuli throughout a session. In fact, some evidence suggests they are extremely reluctant to change their criteria even when different classes of words are marked with different colours (Stretch & Wixted, 1998). Treating each condition separately does not take advantage of these shared criteria and may result in suboptimal model fitting. However, the evidence concerning shared criteria is complex (Benjamin & Bawa, 2004; Dobbins & Kroll, 2005) and does not at this point allow definitive conclusions. In the present study, we present the outcome of both the standard and the aggregated approaches, as they have been described here. This means that we fit a model with shared criteria in the aggregated approach, and use separate criteria in the standard approach. If we can arrive at similar conclusions along these different paths, the assumptions concerning criteria placement are probably not decisive.

Method

Procedure

This experiment was performed as a classroom experiment, in which the stimuli were presented on a large screen, using PowerPoint and a projector, at a rate of 2 s per name. In the ensuing memory test, participants rated names on a 6-point scale, ranging from "I am sure the name was shown" to "I am sure it was not shown before", marking their choices in a booklet. The experiment was divided into two blocks, each containing 64 studied names and a test where they were mixed with 64 distractors. The four types of names were present in equal proportions.

Participants and materials

The experiment took place on two occasions, in classes at two different universities, as part of courses on memory. On the first occasion, 16 participants were tested, and on the second, another 12 (two of which were excluded because of low scores, 65% correct being the cutoff for inclusion). Mean ages were 44 and 27 years, respectively. The stimulus materials were slightly different; the first occasion used an expanded and modified set from

Experiment 1; the second occasion used names drawn from the final set used in Experiments 2 and 4. Despite these slight differences, we analysed the two sets of data together, for simplicity of presentation. Group was included as a between-subjects variable in the analyses, and it turned out that the pattern of effects was quite similar, as shown by the absence of interactions with the Group factor. There were no main or interaction effects pertaining to this factor.

Analysis

Hit rates and false alarm rates were computed to allow comparison with the other experiments. The middle point of the rating scale was used as the criterion.

ROC curves were plotted both on probability axes and on z-transformed axes (see Figure 3). Fitting the ROCs was performed by L. Harvey's program RscorePlus (Harvey, 2005), which uses maximum likelihood estimation.

Results

Variables were submitted to 2 (Frequency) × 2 (Celebrity) × 2 (Group), the latter being a between-participant factor, introduced because tests were performed on two occasions with slightly different materials. It will be mentioned only if significant.

Hit rates were affected by both Celebrity, $F(1, 24) = 67.46$, $p < .001$, $\eta_p^2 = .74$, and Frequency, $F(1, 24) = 17.17$, $p < .001$, $\eta_p^2 = .42$, and there was also an interaction between these two factors, $F(1, 24) = 6.02$, $p = .022$, $\eta_p^2 = .20$.

False alarm rates were affected by Frequency, $F(1, 24) = 13.22$, $p = .001$, $\eta_p^2 = .36$, and Celebrity, $F(1, 24) = 6.47$, $p = .018$, $\eta_p^2 = .21$, and the interaction between them, $F(1, 24) = 4.42$, $p = .046$, $\eta_p^2 = .16$ (Table 4).

Standard ROC analysis. The measure of accuracy for ROCS, d_a, which is similar to d', showed effects of Frequency, $F(1, 24) = 55.84$, $p < .001$, $\eta_p^2 = .70$, and Celebrity, $F(1, 24) = 96.39$, $p < .001$, $\eta_p^2 = .80$, as well as their interaction, $F(1, 24) = 8.50$, $p = .008$, $\eta_p^2 = .26$. The interaction was caused by a larger effect of frequency among nonfamous than among famous names.

The standard deviation for the old distribution, which equals the inverse of the slope of the z-ROC, was tested in a similar ANOVA. It showed a main effect of Celebrity, $F(1, 24) = 4.31$, $p < .049$, $\eta_p^2 = .15$, and no other effect.

Aggregated analysis. Using Matlab's bootstrap function, 10,000 samples of size $n = 26$ were drawn with replacement from the data. Each sample gave

TABLE 4
Experiment 3

	Means					
	Low frequency		*High frequency*		*Effect sizes,* η^2_G	
	Nonfamous	*Famous*	*Nonfamous*	*Famous*	*Frequency*	*Celebrity*
HR	0.69	0.86	0.55	0.83	0.16	0.55
FAR	0.16	0.11	0.27	0.17	0.18	0.14
d_a	1.43	2.20	0.67	1.84	0.39	0.66
SD of old dist.	1.48	1.75	1.42	1.73	0.00	0.06
Mirror effect patterns	HR1	HR2	FAR2	FAR1		
Celebrity	.85	.62	.22	.14		
Frequency	.78	.69	.22	.14		

d_a is a measure of accuracy, the ROC equivalent of d'. *SD* is the standard deviation of the old distribution, i.e., the inverse of slope. ROCs have been fitted to individual data for each condition separately with the maximum likelihood method. For the lower part of the table (the mirror effect patterns), see note to Table 1.

rise to an averaged group ROC, which was submitted to the RScorePlus (Harvey, 2005) program with instructions to fit data for all four types of stimuli together. The result is illustrated in Figure 4.

Of the 10,000 samples, 93% gave acceptable fits ($p > .05$) of the model and were used. From the output of the program, eight means and eight standard deviations (Old/New × High/Low Frequency × High/Low Celebrity) were extracted for each sample and used to compute two contrasts (Low minus High Frequency and High minus Low Celebrity) for both intercepts, $(\mu_{old}—\mu_{new})/\sigma_{old}$, and slopes, $\sigma_{new}/\sigma_{old}$. The vectors of contrast values were sorted and the 2.5 and 97.5 percentiles were identified. The numbers of positive and negative values were tallied, and p_{rep} was computed as the proportion having the same sign as the mean.

Slopes. The Frequency contrast showed an average close to zero: 0.03 (95% CI: −0.31 to 0.37), and a p_{rep} of .57, close to chance. Celebrity, on the other hand, averaged −0.31 (95% CI: −0.63 to 0.04), with a p_{rep} of .96, a reliable effect. Slopes were lower in the High Celebrity conditions.

Intercepts. Both Frequency, $m = 0.89$, $p_{rep} = .999$, and Celebrity, $m = 1.42$, $p_{rep} = 1.00$, had highly replicable effects on intercepts, i.e., on accuracy. This fact confirmed findings from the earlier experiments.

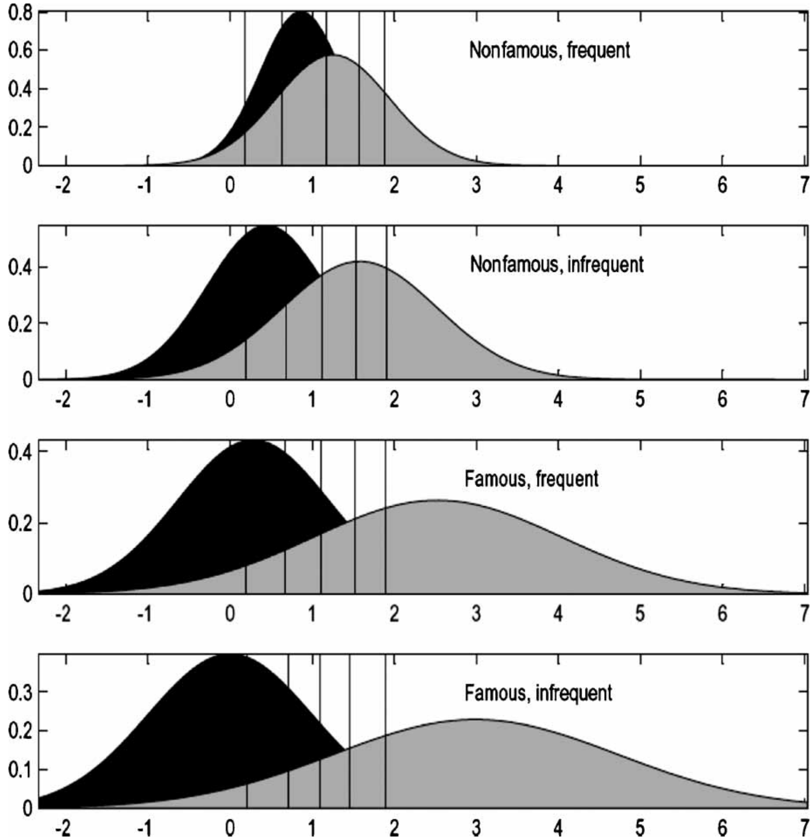


Figure 4. Experiment 3. New (dark) and old distributions of the four simultaneously fitted conditions. Vertical lines mark the criteria.

Discussion

ROC curves traced the same trends concerning accuracy as the earlier experiments, now over a wider range of criteria. Famous names were better retained than nonfamous ones, and infrequent names held an advantage over frequent ones. As before, hit rates were determined by both Celebrity and Frequency, more so by the former. False alarm rates were also influenced by both, but more so by the latter.

The shape and locations of the distributions showed, in general terms, a mirror effect, i.e., those types of stimuli that had great memory strength when they were old, were weak when they were new—i.e., positioned far to the left on the familiarity/memory strength axis. This held true for both famous versus nonfamous names, and for infrequent versus frequent names.

There was a difference between the two factors in their effects on the slopes of the *z*-ROCs. Celebrity affected the slopes, and frequency did not. There has been debate concerning the degree to which slopes vary with accuracy (Heathcote, 2003). One finding has been that variations in materials that affect accuracy also affect slopes. Other, manipulated experimental variables such as study time, often affect accuracy while leaving slopes unchanged.

We found that one aspect of prior experience, the specific knowledge associated with famous names, had an effect on slopes, whereas the nonspecific kind did not. This could be due to the fact that high celebrity raises the level of memory strength—and with it the standard deviation—specifically in the old distribution without affecting the new. High familiarity, on the other hand, raises the familiarity of both the old and the new distributions—and with it the standard deviation—leaving the ratio unaffected.

EXPERIMENT 4

The fourth experiment aimed at reproducing the effects, especially the ROC data, in a new sample of individually tested participants. With the large material divided into shorter blocks, with computer administration and individual testing, we hoped to improve the quality of the ROC data by encouraging use of the full scale of response categories.

Method

Experiment 4 was conducted in the context of ERP (event-related potentials) recording, and electrophysiological data will be reported elsewhere. The focus here is on the behavioural responses and the ROC analyses based on them.

Participants

Twenty-four students at Lund University (14 women) participated in the experiment, which was conducted in a laboratory at the University Hospital. Each participant was tested individually during an approximately 1-hour long session, and received a cinema ticket voucher in compensation. Age of participants averaged 24.8, with a range of 19–42.

Procedure

The pool of 288 names provided stimuli for four blocks, each with 36 studied names and 36 distractors. The four study–test blocks were run consecutively with subject-terminated pauses between them. Assignment of names to study or distractor status was made randomly for each participant, subject to the constraint that each of the four types of names be represented equally in every block.

Stimuli were presented on a computer screen by an E-prime program. Each name was displayed for 2 s in the study phase, preceded by a 1 s fixation cross. In the test phase, a name was first shown for 2 s while responding was disabled (to give ERPs without motor artefacts). Then a prompt appeared above the name ("Have you seen this name before in the experiment?"), and two response buttons appeared below it, marked "Yes" and "No". A mouse click on either response button terminated this display, which was followed by a new prompt ("How sure are you?") with three response buttons, marked "Quite sure", "Relatively sure", and "Not sure". A maximum of 7 s was allowed for response selection, but a response terminated the display, usually much sooner.

Analysis

A 6-point scale, ranging from "Quite sure new" to "Quite sure old" was constructed from the responses and used to plot the ROC and *z*-ROC graphs in Figure 5. Hit rates and false alarm rates were computed, and the ROC data were further analysed.

Harvey's RScorePlus (Harvey, 2005) extracted the parameters d_a and *SD* for each condition and each participant in the standard analysis. Further, using the aggregated scores of the whole group, it fitted all four conditions simultaneously, resulting in the data on which Figure 6 is based. Bootstrapping produced 10,000 samples, which were analysed as in the previous experiment.

Results

Hit rates were affected by Frequency, $F(1, 23) = 32.82$, $p < .001$, $\eta_p^2 = .59$, and Celebrity, $F(1, 23) = 62.64$, $p < .001$, $\eta_p^2 = .73$, and their interaction, $F(1, 23) = 18.47$, $p < .001$, $\eta_p^2 = .45$. False alarm rates were affected by Frequency, $F(1, 23) = 27.51$, $p < .001$, $\eta_p^2 = .55$, and Celebrity, $F(1, 24) = 10.36$, $p = .004$, $\eta_p^2 = .31$, with no interaction (Table 5).

Standard analysis. The measure of accuracy for ROCs, d_a, the analogue of d', showed effects of Frequency, $F(1, 23) = 46.41$, $p < .001$, $\eta_p^2 = .67$, and Celebrity, $F(1, 23) = 81.21$, $p < .001$, $\eta_p^2 = .78$, as well as their interaction, $F(1, 23) = 5.09$, $p = .034$, $\eta_p^2 = .18$. As before, famous and infrequent names

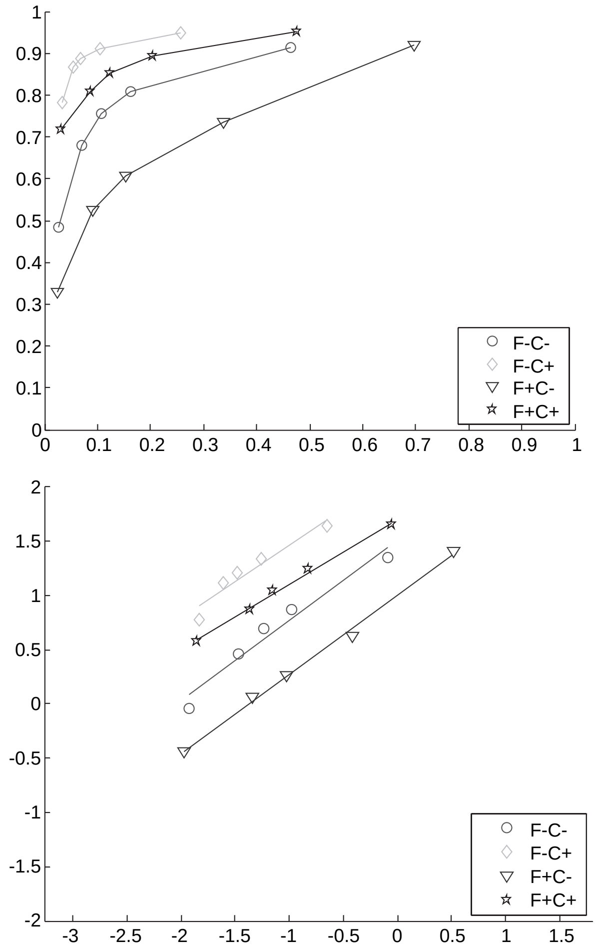


Figure 5. ROC curves (upper half) and z-transformed ROC-curves (lower half) for Experiment 4. Markers: circles: low frequency, low celebrity; diamonds: low frequency, high celebrity; triangles: high frequency, low celebrity; stars: high frequency, high celebrity.

were better remembered than their nonfamous and frequent counterparts. The interaction indicated a larger frequency effect among the nonfamous.

Standard deviation of the old distribution, the inverse of slope of the z-ROCs, showed only an effect of Celebrity, $F(1, 23) = 5.26$, $p = .031$, $\eta_p^2 = .19$, all other Fs < 1.

Aggregated. Of 10,000 bootstrapped samples, 94% gave acceptable fits. Contrasts for high-low celebrity and low-high frequency were computed, the order of the terms arranged to place higher performance first.

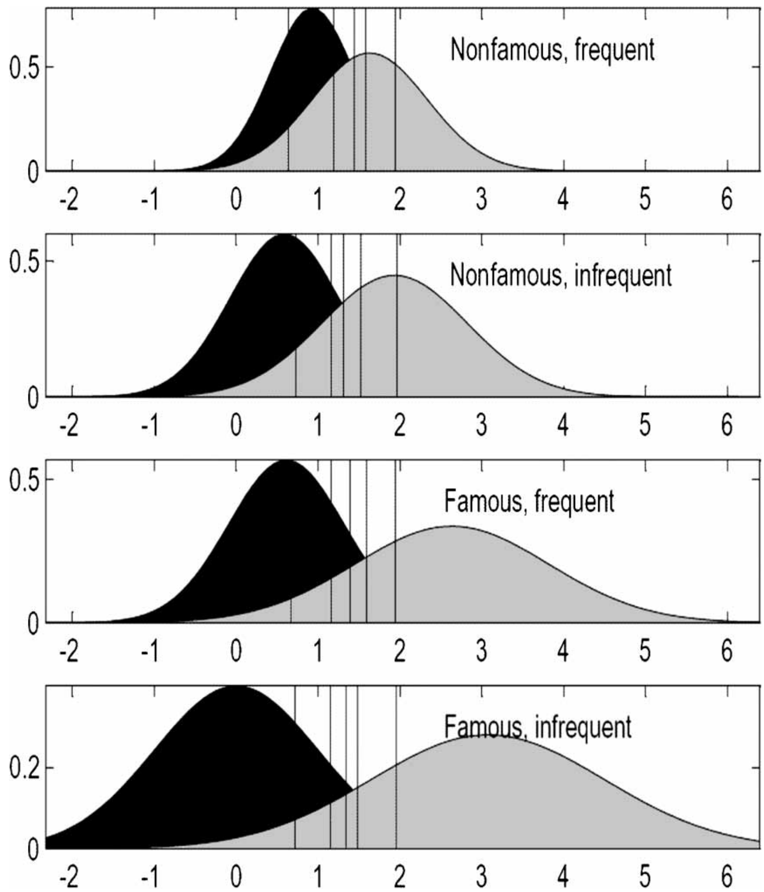


Figure 6. Experiment 4. New (dark) and old distributions of the four simultaneously fitted conditions. Vertical lines mark the criteria.

Slopes. The replicability was only marginal, $p_{rep} = .86$ for both Frequency and Celebrity, whereas .90 could probably be considered the lower limit for a significant effect. The direction of the effect was such that high celebrity produced lower slopes (and higher accuracy). The direction of the Frequency effect, on the other hand, was such that low frequency produced *higher* slopes (and higher accuracy). The means were $+0.17$ (frequency) and -0.18 (celebrity).

Intercepts. Both contrasts evidenced highly reproducible effects; both $p_{rep} = 1.00$. Means of the contrasts were 0.98 and 1.32, for Frequency and Celebrity, respectively.

TABLE 5
Data from Experiment 4

	Means					
	Low frequency		*High frequency*		*Effect sizes, η^2_G*	
	Nonfamous	*Famous*	*Nonfamous*	*Famous*	*Frequency*	*Celebrity*
HR	0.76	0.89	0.61	0.85	0.27	0.60
FAR	0.11	0.07	0.15	0.12	0.25	0.13
d_a	1.78	2.50	1.18	2.18	0.34	0.65
SD of old distribution	1.37	1.57	1.38	1.84	0.01	0.05
Mirror effect patterns	HR1	HR2	FAR2	FAR1		
Celebrity	.87	.69	.13	.10		
Frequency	.83	.73	.14	.09		

For the lower part of the table, see note to Table 1.

Item analysis. In a regression analysis over items, hit rate was entered as the dependent variable, and the potential predictors—(a) the log-transformed number of hits in the Internet search of the telephone directory (*freq_hits*), and (b) the log-transformed number of hits in the Internet search of the media sites (*celeb_hits*)—were introduced in a stepwise analysis. Both predictors were accepted as significant, c*eleb_hits*: $\beta = .50$, $t(285) = 10.24$, $p < .001$; *freq_hits*: $\beta = -.30$, $t(285) = -6.27$, $p < .001$.

In a similar analysis, using false alarm rate as the dependent variable, frequency was found to be a significant predictor, *freq_hits*: $\beta = .29$, $t(285) = 5.10$, $p < .001$, and so was c*eleb_hits*: $\beta = -.14$, $t(285) = -2.43$, $p = .016$.

Discussion

As in all the other experiments, we found a partial dissociation, in that frequency affected false alarm rates more than celebrity did, and celebrity affected hit rates more than frequency did. This held true for both analysis over subjects and over items.

Thus, although both factors had large effects on net accuracy (d_a), the patterns of effects on the components of performance were different. As to the ROC data, celebrity had an impact on the standard deviations of the distributions in the standard analysis, whereas frequency had none. (In the aggregated analysis, both had only weak effects, however, in opposite directions.) We interpret this as indicating that high frequency impairs performance by raising the level—and importantly, the variance—of familiarity in both old and new distributions.

In other words, we find our results to be compatible with a pattern where frequency primarily raises the variance of old and new distributions alike, and impairs performance as a result. Celebrity, on the other hand, raises mean and variance of the old distribution specifically, improving performance in the process.

GENERAL DISCUSSION

We studied the effects of two types of prior experience on the memorability of names. Both were environmental variables, presumably related to experiential frequency, such that higher values on the two variables meant higher probabilities of encountering the names in daily life, given the relatively homogeneous cultural environment of our participants. The variables—name frequency and celebrity—were measured both by environmental statistics and by participant ratings, with satisfactory reliability.

Although similar in their relation to experiential frequency, the variables had completely different effects on memory accuracy. The effects were different not only in size, but more strikingly, in direction. Increasing name frequency lowered accuracy, a pattern reminiscent of the much-studied word frequency effect. Increasing name celebrity, on the other hand, raised accuracy. In this latter respect, our data resembled earlier studies, where knowledge of the stimulus domain has been seen to improve accuracy. Examples include superior memory for chess positions in chess masters (de Groot, Gobet, & Jongman, 1996), and the memorial advantage for words taken from a student's major field of study (Allen & Garton, 1968; Chalmers et al., 1997). Wine experts show superior recognition memory for wine-related odours (Parr, White, & Heatherbell, 2004), and experienced golfers show enhanced memory for specific putts, but only when routinisation of the putting task is disturbed (Beilock, Wierenga, & Carr, 2002).

The divergent effects of prior experience are probably quite general, rather than specific to names. The pattern of sharpened memory within preferred domains of knowledge can arguably be ascribed to increased opportunities for elaborative encoding, although this would need independent evidence. The other pattern is perhaps more counterintuitive, although we can easily find examples in daily life of the familiarity that breeds disregard. In fact, many memory failures are the result of poor encoding of run-of-the-mill events, or neglect by habit. As the history of the word frequency effect shows, this common phenomenon has proven recalcitrant for several memory theories.

Recently, neuroscience has turned up evidence (Fernandez & Tendolkar, 2006) that a part of the medial temporal lobe, the rhinal cortex, acts as a gatekeeper to the elaborated encoding orchestrated by the hippocampus.

If familiarity is high, as judged by the rhinal cortex, the event is deemed uninteresting and it is denied access to deepened encoding. If news value is high, on the other hand, all the facilities of the memory system are called upon to engrave the new event. In a recent fMRI study, stimuli that had been primed beforehand, underwent less encoding and showed depressed retention in relation to novel stimuli (Wagner, Maril, & Schacter, 2000).

As the present data show, some well-known objects escape the ban of the gatekeeper. If a name is famous, be it ever so common, it enjoys the privilege of the newsworthy, and gets to be encoded deeply. What determines whether an object, although familiar, can pass the scrutiny of the gatekeeper? We have proposed that specific semantic knowledge, i.e., the fact that it is individuated, is decisive. If a name is ripe with unique and detailed associations, a web of potential retrieval cues can be established at encoding, possibly bound together by hippocampal pointers. If, in addition, rhinal cortex judges the item to be novel, this seal of approval further facilitates encoding. The fact that we found fame and novelty to interact overadditively in Experiments 1 and 2 suggests that they both affect the same processing stage, possibly the hippocampal encoding mechanism.

Apart from the difference in direction of the general memory effect, we also found that the two types of prior experience had different profiles in their impact on memory components. This can be summarised in two points:

- Celebrity influenced hit rates more than frequency did, and frequency influenced false alarm rates more than celebrity did.
- Celebrity tended to have some effect on ROC slopes, but frequency did not.

The first of these points was evident in the measures of effect size, which showed consistency across the four experiments. However, to test the statistical significance of this fact further, we formed the contrast of high versus low celebrity, and low versus high frequency and computed these contrasts over hit rates and false alarm rates for each individual. The outcome is shown in Figure 7. Furthermore, 2×2 ANOVAs were computed with these contrast scores, and all four experiments evidenced the critical effect, a significant interaction (all ps at least $< .02$), showing the celebrity contrast being larger for hit rates, and the frequency contrast being larger for false alarm rates. The mean replicability of this effect, p_{rep} (Killeen, 2005), is better than .97. The fact that hit rates and false alarm rates can be dissociated has been interpreted as evidence in favour of two-process theories of recognition memory (Reder et al., 2000).

In the examination of ROCs, we found indications that celebrity affected slopes, but frequency did not. Earlier literature has discussed why some variables (especially those related to materials) increase both accuracy and

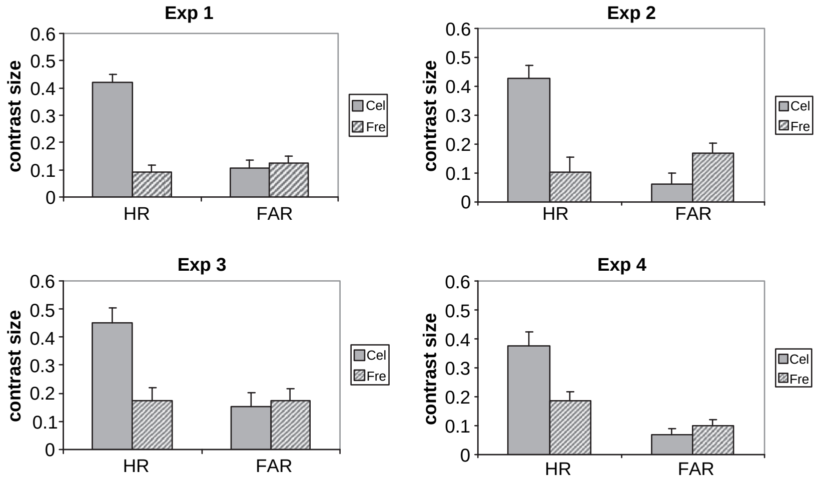


Figure 7. Means of the contrasts high–low celebrity and high–low frequency, computed for hit rates and false alarm rates in the four experiments.

ROC slopes, whereas others affect accuracy without changing slopes, and no definitive consensus has been reached (Heathcote, 2003). In any event, the finding adds to the evidence that the two variables exert different effects.

The frequency effect can tentatively be characterised as belonging to semantic memory, or possibly to a very long-term form of conceptual priming. This is not denying the fact that there are related phenomena in the animal learning literature. Conditioning can be impeded by *Kamin blocking*. This refers to the finding that a stimulus can be rendered ineffective as elicitor of a conditioned response with which it is paired, if it (the stimulus) has been familiarised beforehand. In anthropomorphic terms, the animal discounts the stimulus as being useless as a predictor, because of its prior experience with it in a noncausal role. A very similar phenomenon, *latent inhibition*, has been examined in some human learning studies (Lubow & Gewirtz, 1995) that have demonstrated how stimuli can be rendered ineffective for learning (i.e. conditioning) by frequent presentations before conditioning starts, evidently a case of the familiarity that breeds discounting.

A recent theory has awarded familiarity a third place in the memory hierarchy, subordinate to declarative memory, but on an equal footing with semantic and episodic memory (Moscovitch et al., 2005), and Fernandez and Tendolkar (2006) have reviewed the evidence for the rhinal cortex acting as a familiarity detector. The phenomena of familiarity versus novelty are evidently very important in regulating encoding resources up and down in the course of everyday experience, thereby deciding what becomes preserved in memory and what disappears without a trace.

In summary, we have found indications of two types of memory, tentatively described as semantic. They are both related to the prior experience of names, but they exert diametrically opposed effects on episodic recognition. One type is related to general familiarity. At the time of study, it tends to inhibit deep encoding by signalling ordinariness, thereby denying access to encoding resources. At the time of retrieval, it confuses the recogniser by engendering a sense of familiarity that is hard to distinguish from the sought-after memory. By this mechanism, the false alarm rate is raised. Further, it raises the level of variability in both the old and the new distributions.

The distinctive type of semantic memory, on the other hand, resembles knowledge of subject matter, or expertise within a stimulus domain. It facilitates encoding by providing links to individuating features, highly specific associations that serve to make the memory distinctive. This effect on encoding raises hit rates without notably affecting false alarm rates.

Original manuscript received December 2006
Revised manuscript received March 2007
First published online July 2007

REFERENCES

Allen, L. R., & Garton, R. F. (1968). the influence of word-knowledge on the word-frequency effect in recognition memory. *Psychonomic Science*, *10*(12), 401–402.

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, *2*(6), 396–408.

Anderson, J. R., & Schooler, L. J. (2000). The adaptive nature of memory. In E. Tulving & F. Craik (Eds.), *The Oxford handbook of memory.* Oxford, UK: Oxford University Press.

Arndt, J., & Reder, L. M. (2002). Word frequency and receiver operating characteristic curves in recognition memory: Evidence for a dual-process interpretation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(5), 830–842.

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, *37*(3), 379–384.

Bartlett, F. (1932). *Remembering*. Oxford, UK: Oxford University Press.

Beilock, S. L., Wierenga, S. A., & Carr, T. H. (2002). Expertise, attention, and memory in sensorimotor skill execution: Impact of novel task constraints on dual-task performance and episodic memory. *Quarterly Journal of Experimental Psychology: Section A*, *55*(4), 1211–1241.

Benjamin, A. S., & Bawa, S. (2004). Distractor plausibility and criterion placement in recognition. *Journal of Memory and Language*, *51*(2), 159–172.

Chalmers, K. A., & Humphreys, M. S. (1998). Role of generalized and episode specific memories in the word frequency effect in recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(3), 610–632.

Chalmers, K. A., Humphreys, M. S., & Dennis, S. (1997). A naturalistic study of the word frequency effect in episodic recognition. *Memory and Cognition*, *25*(6), 780–784.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behaviour*, *12*, 335–359.

Clark, S. E., & Gronlund, S. D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin and Review*, *3*(1), 37–60.

De Groot, A. D., Gobet, F., & Jongman, R. W. (1996). *Perception and memory in chess: Studies in the heuristics of the professional eye*. Assen, The Netherlands: Van Gorcum & Co.

Dennis, S. (1995). The impact of the environment on the word frequency and the null list strength effects in recognition memory [Electronic version]. *Noetica, 1*. Retrieved 19 March 2007 from http://www2.psy.uq.edu.au/CogPsych/Noetica/Issue1/Environment.html

Dobbins, I. G., & Kroll, N. E. A. (2005). Distinctiveness and the recognition mirror effect: Evidence for an item-based criterion placement heuristic. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(6), 1186–1198.

Estes, W. K., & Maddox, W. T. (2002). On the processes underlying stimulus-familiarity effects in recognition of words and nonwords. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(6), 1003–1018.

Fernandez, G., & Tendolkar, I. (2006). The rhinal cortex: "Gatekeeper" of the declarative memory system. *Trends in Cognitive Sciences*, *10*(8), 358–362.

Glanzer, M., & Adams, J. K. (1985). The mirror effect in recognition memory. *Memory and Cognition*, *13*(1), 8–20.

Glanzer, M., & Adams, J. K. (1990). The mirror effect in recognition memory: Data and theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(1), 5–16.

Glanzer, M., Adams, J. K., Iverson, G. J., & Kim, K. (1993). The regularities of recognition memory. *Psychological Review*, *100*(3), 546–567.

Glanzer, M., Kim, K., Hilford, A., & Adams, J. K. (1999). Slope of the receiver-operating characteristic in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(2), 500–513.

Harvey, L. O. (2005). RscorePlus: A Macintosh/Windows program that fits a multiple-distribution, variable-criterion signal detection theory model to confidence-rating data (Ver. 5) [Computer program]. http://psych.colorado.edu/ ~ lharvey/Software%20Zip%20Files/RscorePlus_Win.zip

Heathcote, A. (2003). Item recognition memory and the receiver operating characteristic. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(6), 1210–1230.

Hirshman, E., & Henzler, A. (1998). The role of decision processes in conscious recollection. *Psychological Science*, *9*(1), 61–65.

Jacoby, L. L., Jones, T. C., & Dolan, P. O. (1998). Two effects of repetition: Support for a dual-process model of knowledge judgments and exclusion errors. *Psychonomic Bulletin and Review*, *5*(4), 705–709.

Kelley, C. M., & Jacoby, L. L. (1998). Subjective reports and process dissociation: Fluency, knowing, and feeling. *Acta Psychologica*, *98*(2–3), 127–140.

Killeen, P. R. (2005). An alternative to null-hypothesis significance tests. *Psychological Science*, *16*(5), 345–354.

Lubow, R. E., & Gewirtz, J. C. (1995). Latent inhibition in humans: Data, theory, and implications for schizophrenia. *Psychological Bulletin*, *117*(1), 87–103.

Maddox, W. T., & Estes, W. K. (1997). Direct and indirect stimulus-frequency effects in recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(3), 539–559.

Martinez, W. L., & Martinez, A. R. (2002). *Computational statistics handbook with Matlab*. London: Chapman Hall.

Moscovitch, M., Rosenbaum, R. S., Gilboa, A., Addis, D. R., Westmacott, R., Grady, C., et al. (2005). Functional neuroanatomy of remote episodic, semantic and spatial memory: A unified account based on multiple trace theory. *Journal of Anatomy*, *207*(1), 35–66.

Parr, W. V., White, K. G., & Heatherbell, D. A. (2004). Exploring the nature of wine expertise: What underlies wine experts' olfactory recognition memory advantage? *Food Quality and Preference*, *15*(5), 411–421.

Raaijmakers, J. G. W., Schrijnemakers, J. M. C., & Gremmen, F. (1999). How to deal with "the language-as-fixed-effect fallacy": Common misconceptions and alternative solutions. *Journal of Memory and Language*, *41*(3), 416–426.

Rajaram, S. (1996). Perceptual effects on remembering: Recollective processes in picture recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(2), 365–377.

Ratcliff, R., McKoon, G., & Tindall, M. (1994). Empirical generality of data from recognition memory receiver-operating characteristic functions and implications for the global memory models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 763–785.

Reder, L. M., Nhouyvanisvong, A., Schunn, C. D., Ayers, M. S., Angstadt, P., & Hiraki, K. (2000). A mechanistic account of the mirror effect for word frequency: A computational model of remember-know judgments in a continuous recognition paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(2), 294–320.

Schooler, L. J., & Shiffrin, R. M. (2005). Efficiently measuring recognition performance with sparse data. *Behavior Research Methods*, *37*(1), 3–10.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, *117*(1), 34–50.

Stretch, V., & Wixted, J. T. (1998). Decision rules for recognition memory confidence judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(6), 1397–1410.

Wagner, A. D., Maril, A., & Schacter, D. L. (2000). Interactions between forms of memory: When priming hinders new episodic learning. *Journal of Cognitive Neuroscience*, *12*(6), 52–61.

Westmacott, R., & Moscovitch, M. (2003). The contribution of autobiographical significance to semantic memory. *Memory and Cognition*, *31*(5), 761–774.

Wixted, J. T., & Stretch, V. (2004). In defense of the signal detection interpretation of remember/know judgments. *Psychonomic Bulletin and Review*, *11*(4), 616–641.

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.

Paper II

Familiarity or Conceptual Priming: Event-related Potentials in Name Recognition

Georg Stenberg[1], Johan Hellman[2], Mikael Johansson[2], and Ingmar Rosén[2]

Abstract

■ Recent interest has been drawn to the separate components of recognition memory, as studied by event-related potentials (ERPs). In ERPs, recollection is usually accompanied by a late, parietal positive deflection. An earlier, frontal component has been suggested to be a counterpart, accompanying recognition by familiarity. However, this component, the FN400, has alternatively been suggested to reflect a form of implicit memory, conceptual priming. The present study examined the ERP components of recognition memory using an episodic memory task with a stimulus material consisting of names, half of which were famous. Along a different dimension, the names varied in how rare or common they were. These dimensions, frequency and fame, exerted powerful effects on memory accuracy, and dissociated the two recognition processes, such that frequency gave rise to familiarity and fame fostered recollection, when the receiver operating characteristics data were analyzed with Yonelinas' dual-process signal detection model. The ERPs corresponded fully to the behavioral data because frequency affected the frontal component exclusively, and fame affected the parietal component exclusively. Moreover, a separate behavioral experiment showed that conceptual priming was sensitive to fame, but not to frequency. Our data therefore indicate that the FN400 varies jointly with familiarity, but independently of conceptual priming. ■

INTRODUCTION

Recognition is a deceptively simple form of memory. In the layman's view, it is a matter of comparing the present stimulus to stored images in memory and finding a match. Yet we frequently come across the experience of knowing a match is there, despite having searched memory unsuccessfully. A face can look hauntingly familiar, but defy all our attempts to match it with a name. Experiences such as these have inspired the view that recognition consists of two processes, in at least partial independence of each other: familiarity and recollection. We can have one without the other, such as when we fail to recollect where we met the person with the familiar face (Mandler, 1980).

In explaining this phenomenon, researchers have chosen two different paths (Yonelinas, 2002). Dual-process theories have proposed that recognition makes use of two qualitatively different types of information. Familiarity is based on a nonspecific computation of memory strength, subjectively accessible as a general feeling of knowing. Recollection, on the other hand, makes use of more detailed information, is often specific about time and place, and helps reconstruct the episode when the stimulus was encountered.

Single-process theories, in contrast, do not deny that recognition can give rise to a spectrum of experiences, from the vaguest to the most detailed, but claim that they can all be placed on the same continuum. Thus, no separate kind of information is accessed to dignify a vague feeling of familiarity into full-blown recollection, only more of the same, held with more confidence. Single-process theories have often been articulated using the concepts of signal detection theory, and quantitative models (Clark & Gronlund, 1996) have been remarkably successful in explaining a variety of memory phenomena within this framework. This parsimony is not given up lightly, and the burden of proof rests upon dual-process theories to justify claims to the contrary.

Familiarity is assumed to operate quickly, giving a general estimate of the similarity between an item and the contents of memory (Hintzman & Curran, 1994). Recollection is the more effortful of the two. If successful, it gives access to a stored representation of an earlier event in its full context, allowing a re-experience of the episode. It is often thought that familiarity is a graded and stochastic process, well described by signal detection theory. More disputed is the contrasting claim that recollection is an all-or-none phenomenon, described by high-threshold theory (Parks & Yonelinas, 2007; Wixted, 2007).

Ways of separating the two recognition processes using behavioral methods have been developed. The

[1]Kristianstad University, Sweden, [2]Lund University, Sweden

remember/know procedure lets the participant characterize her subjective experience by prompting her with a follow-up question after each affirmative recognition decision: ''Do you remember seeing the word, or do you just know it was there?''(Gardiner & Richardson-Klavehn, 2000) Another method takes advantage of the observation that recollection almost always is accompanied by a high degree of confidence, whereas familiarity can run the whole gamut, from conviction to guessing. The method, which has been devised by Andrew Yonelinas (2001a), is based on data gathered in the receiver operating characteristics (ROC) paradigm, that is, a recognition task where decisions are made on a (usually) 6-point scale, running from ''Quite sure it is new'' to ''Quite sure it is old.'' A model is fitted to the accumulated data, in which a familiarity component accounts for the signal detection share of the responses, and a recollection component accounts for the affirmative responses at the high-confidence end of the scale.

Experimental manipulations, such as a source monitoring task, can expose the kind of contextual knowledge that recollection is privy to. Hence, conditions in which source judgments are correct can be assumed to rely on recollection. Further, manipulations of processing levels at encoding can produce the desired type of memory (Rugg et al., 1998): Deep encoding makes recollection more likely, and shallow encoding makes it unlikely, leaving familiarity as the only resort.

The great interest in the two processes of recognition has been staked on the claim that different neural structures are involved. Based on extensive experimentation with animal analogues of familiarity and recollection, key roles have been assigned to the perirhinal cortex for the former and to the hippocampus for the latter (Aggleton & Brown, 2006; Brown & Aggleton, 2001).

The human neuroimaging and neuropsychology data have recently been reviewed (Skinner & Fernandes, 2007). The relevant fMRI studies show activation patterns compatible with, but not restricted to, the expected areas of the medial-temporal lobe. Areas of prefrontal and parietal cortices as well as content-specific sensory regions are also brought into play by recognition processes. Recollection produces activation that it shares with familiarity, and moreover, brings additional regions into a more extensive and possibly a more coherent network (Skinner & Fernandes, 2007). The clinical data show areas in which lesions harm recollection but not familiarity (Yonelinas et al., 2004), but the opposite pattern has been harder to find. Recently, however, a double dissociation has been discovered, in which the volume of the hippocampus correlated with recollection in a group of elderly people, and, importantly, the volume of the entorhinal cortex correlated with familiarity (Yonelinas et al., 2007).

Event-related potentials (ERPs) have provided converging evidence (recently reviewed by Rugg & Curran, 2007). The proposed two processes of recognition have been associated with two characteristic electrophysiological signatures, which can be seen as components in old–new difference waveforms. The slower process of recollection gives rise to a relatively late (400–900 msec) component (Rugg & Yonelinas, 2003) with a parietal distribution, often with a left-over-right asymmetry. Behavioral evidence gathered jointly with ERPs testify in favor of the connection with recollection. Thus, the component is sensitive to levels-of-processing manipulations (Paller, Kutas, & McIsaac, 1995), and larger for items eliciting ''remember'' rather than ''know'' responses (Curran, 2004). It responds to source memory tasks and associative recognition even more than to item memory. It is drastically curtailed by lesions to major memory areas in the brain such as the medial-temporal lobe (Duzel, Vargha-Khadem, Heinze, & Mishkin, 2001). From what is known through other methods, there are reasons to believe that the hippocampus takes a crucial part, through interactions with the cortex, in the encoding and retrieval of recollected memories (O'Reilly & Norman, 2002). The parietal scalp distribution is thought to reflect mainly the cortical contributions to this exchange.

Following the parietal old–new positive deflection, one additional component may appear under some conditions. It is a late posterior negativity (Johansson & Mecklinger, 2003), which appears with an onset latency of approximately 800 msec and a mainly parietal distribution. It is elicited in relatively demanding tasks where the participant is expected to determine in what context the stimulus was encountered (e.g., Johansson, Stenberg, Lindgren, & Rosen, 2002; Cycowicz, Friedman, & Snodgrass, 2001) or is otherwise made to monitor his actions closely for errors (Curran, DeBuse, & Leynes, 2007; Herron, 2007; Nessler & Mecklinger, 2003).

The large parietal component was apparent in early stages of research into old–new ERP effects. A component corresponding to familiarity is a more recent finding (Mecklinger, 2006; Curran, 2000). It is called the FN400, or the mid-frontal old–new effect, because it is seen as a frontal, positive-going modulation of a negative component with a peak around 400 msec. A number of characteristics have made an association with familiarity seem likely. It is elicited by lures that are closely similar to studied items, such as the word ''horses'' when ''horse'' has been studied, or a mirror-reversed version of a studied image (Mecklinger, 2006; Curran & Cleary, 2003; Curran, 2000). Unlike the parietal effect, the FN400 seems sensitive only to the general similarity, not to the finer distinctions between targets and lures, even when instructions emphasize them. Its early latency (300–500 msec) is compatible with a fast-acting process. Further, it covaries with response confidence, as would be expected of a continuous signal-detection process (Woodruff, Hayama, & Rugg, 2006).

However, another interpretation has been proposed, which associates the FN400 not with familiarity, but with conceptual priming (Paller, Voss, & Boehm, 2007). Paller, Voss, and Boehm (2007) observed that many of the tasks in which the FN400 has been produced contain an element of conceptual overlap between study and test. For example, the lures with altered plurality in Curran's (2000) study were probably conceptually primed in the study phase to roughly the same degree as preserved-plurality targets. This priming could produce the attenuation of an N400 that has become known as the FN400. Supporting this alternative explanation, the authors adduced studies with negative findings where the familiarity hypothesis would have predicted an FN400 effect. A case in point comes from a study (Yovel & Paller, 2004) of face recognition with a contextual memory component. The task was to learn an occupation presented auditorily along with each face. Faces that were recognized without retrieval of the occupation were said to be recognized by familiarity alone. These faces produced no FN400, only the parietal old–new effect, but with lower amplitude than the recollected face–occupation pairs. Instead, in a study where conceptual priming was measured, it was found to correlate with a frontal old–new effect, whereas a measure of explicit memory correlated with a parietal effect (Voss & Paller, 2006).

A prediction can be deduced from Paller et al.'s position, namely, that items without conceptual content produce no FN400, no matter how familiar they may be. The prediction applies to pseudowords and nonfigurative drawings, among other things. Unfortunately, the data hitherto are not very clear on this point. Completely meaningless stimuli are hard to find, considering that meaning is in the eye of the beholder. When pseudowords and quasi-drawings do give rise to FN400s and have been claimed as evidence against the conceptual-priming interpretation, their meaninglessness has been contested (Paller et al., 2007; Rugg & Curran, 2007).

Purpose

Our aim was to dissociate recollection from familiarity behaviorally to permit study of the electrophysiological signatures of the two processes in retrieval. As a tool, we use a stimulus set of names that we have used before for the same purpose (Stenberg, Hellman, & Johansson, 2008). The idea is to let participants' previous (pre-experimental) acquaintance with the material determine the way in which it will be remembered. The material varies along two dimensions that we have found to predispose toward familiarity and recollection. We will verify that they do so in the present sample also, using Yonelinas' method of analyzing ROC data, the dual-process signal detection (DPSD) model.

The stimulus material consists of names, a common object for semantic memory (and with age, a proverbial source of memory lapses). The set of names we used (first name plus last name, all Swedish) vary in media exposure, some of them being names of well-known celebrities with frequent appearances in news media, others being completely incognito. Along a different dimension, they also vary in how rare or common they are, some being shared by many namesakes, others being unique. Examples of famous, rare names (in English-speaking countries) would be Barack Obama and Gwyneth Paltrow. Common, famous names are, for example, Jessica Simpson and Will Smith, and common nonfamous names are Tom Williams and Jane Wilson. Rare, nonfamous names would be, for example, Sebastian Weisdorf and Guido Bagnaschi.

Elsewhere, we have proposed that these two dimensions correspond to two types of semantic memory (Stenberg et al., 2008). In the present context, we are concerned with how they interact with episodic memory, and in particular, we suggest a special relation with the dual processes of recollection and familiarity.

Famous names are prone to be recalled with a rich context of remembered associations. Because of the prior knowledge we possess about the person who bears it, the name is readily woven into a web of associated facts, thoughts, and images that can serve as aids in future retrieval attempts. This aspect of the memory for famous names lets it provide material for recollection.

Familiarity, the other major process of recognition memory, is what we suggest to be related to name frequency, that is, the probability of coming across the name in daily life. A common name, such as John Smith, will, no doubt, sound familiar on first presentation. But it will not be easily remembered, being subject to interference from many other, similar instances. The sense in which we suggest that name frequency operates on memory by familiarity is quite the opposite: *Uncommon* names are recognized by the sense of familiarity they evoke. This conjecture, the familiarity increment hypothesis (Mandler, 1980), will be elaborated a little further ahead.

In brief, we expect a relation to hold between properties of our stimulus material and the processes they set in motion in a memory experiment. The dimensions of celebrity and frequency are varied orthogonally in our set of names, whose validity has been checked by searches of the Internet. We counted the number of times each name was used on national mass media Web sites; we included a sample of morning newspapers, tabloids, and television networks. The number served as a measure of celebrity. Acting as proxy for name frequency was the number of persons bearing the name listed in a national, Web-accessible telephone directory. In an earlier study (Stenberg et al., 2008), we have validated these qualities in our name database, showing good agreement between participant ratings and the Internet search statistics. And crucially, we demonstrated strong, independent effects of both variables on episodic recognition performance. Famous names were better remembered than nonfamous ones, infrequent names better than frequent ones. To these findings we now wish to add an examination of the brain processes involved. (Some aspects of the behavioral data from the present experiment were reported as Experiment 4 in Stenberg et al., 2008).

Hypotheses

We expect application of the Yonelinas model to reveal the extent to which different categories of names are recognized by recollection or familiarity. We expect a mapping between recollection and celebrity in our material, such that famous names are recollected more often than nonfamous ones. Similarly, we expect a mapping between familiarity and frequency, such that infrequent names seem more familiar at the time of test than frequent ones (in a sense to be explained shortly).

Yonelinas's model has been deployed in studies of ordinary item recognition as well as of associative recognition (Yonelinas, 1997), of source memory (Yonelinas, 1999), and of amnesic patients (Yonelinas, Kroll, Dobbins, Lazzara, & Knight, 1998). It has been validated against other methods of separating dual processes, such as Jacoby's process dissociation procedure (Yonelinas, 1994) and the remember/know procedure (Yonelinas, 2001b). However, to our knowledge, the model has not been applied previously in direct connection with ERP measures of recollection and familiarity.

Our expected mapping between recollection and celebrity, on the one hand, and familiarity and frequency, on the other, may turn out to involve full, partial, or no overlap between the two processes. As a working hypothesis, we choose the simpler alternative and assume a double dissociation. In other words, we assume a relation between recollection and celebrity alone, on the one hand, and familiarity and frequency alone, on the other.

The Familiarity Increment Hypothesis

The ways in which accuracy in a memory test is affected by a name's frequency deserve working out in some detail. We will use the term memory strength to denote the dimension underlying memory decisions in global matching models of memory (Murdock, 1993; Gillund & Shiffrin, 1984; Hintzman, 1984). Memory strength arises as a function of comparisons performed against all stored memory items. Therefore, obviously, an *infrequent* name has lower memory strength at the outset, that is, when presented for the first time in the experiment than a frequent name. In the test phase, a studied infrequent name has gained something in strength as a result of the study episode, but it may still be lower in strength than a studied, or even an unstudied, frequent name. However, the testee is likely to make his decision based on a Bayesian mechanism, taking into account the prior memory strength, given the frequency of the name. In other words, a rare name has low a priori memory strength, and in view of this, the participant may very well decide that a tiny elevation of strength is a telltale sign of its being studied. In reference to the word frequency effect on recognition, Mandler (1980) wrote in his seminal paper (p. 267): "On the basis of familiarity alone, the best explanatory candidate is an incremental explanation; that is, the additional presentation produces a larger relative increment for low than for high frequency words." This view is also in agreement with the BIC (binding of items and context) model of medial-temporal lobe functioning (Diana, Yonelinas, & Ranganath, 2007) and the Gatekeeper model (Fernandez & Tendolkar, 2006), according to which activity in the perirhinal cortex is high when encoding new and unfamiliar objects, and correspondingly low when recognizing the familiar.

This Bayesian likelihood mechanism is the mainspring in most modern explanations of the mirror effect. In the present context, it provides the explanation for better recognition of rare than of common names. For example, in the memory model REM (Shiffrin & Steyvers, 1997), likelihood ratios are formed for each feature and combined into overall odds for the item being old versus the item being new. In this process, unusual features contribute heavily because of their diagnosticity. The probability of an unusual name matching even a faint memory trace by chance is low, and the memory mechanism takes this into account automatically when deciding that the item is old. It is worth emphasizing that this Bayesian weighing of the evidence takes place for each and every item, as part of the standard processing of stimuli. Thus, it is not a recourse of extra deliberation, consciously chosen to resolve a temporarily suspended judgment on a thorny issue. The mechanism deciding about memory strength acts quickly and automatically, corresponding to the familiarity component of two-process theories. And that is the reason we expect the infrequent names to be better recognized than frequent names, by virtue of their familiarity, paradoxical as this may sound.

ERP Hypotheses

In the test phase ERPs, we anticipate old–new differences in the early stages (300–500 msec) as a positive deflection over frontal areas, the FN400 effect. This effect, we hypothesize, will be greater for the types of stimuli that are primarily recognized on the basis of familiarity (i.e., the infrequent names).

We also wish to compare the conditions in which we find an FN400 with the conditions that produce conceptual priming. If the conceptual priming hypothesis for the FN400 is tenable, conditions conducive to priming should also produce the FN400. If, on the other hand, the familiarity hypothesis is right, conditions that give behavioral signs of recognition by familiarity should also produce the FN400.

In a later stage (400–900), we expect to see large old–new differences as a positive deflection, maximal over parietal areas. If our hypothesis is correct, it will be greater for famous names than for nonfamous ones because the former can be recollected more readily.

A late posterior negativity can be expected for names that are ambiguous as to source. In our material, many of the names are well known pre-experimentally, particularly the names of celebrities, but also the more common names. These stimuli necessitate close scrutiny to determine whether the name was recently seen in the media or in the experiment. This is the type of task that has been known to produce the late posterior negativity (Johansson & Mecklinger, 2003).

EXPERIMENT 1

Methods

Participants

Twenty-four students at Lund University (14 women) participated in the first experiment, which was conducted in a laboratory at the University Hospital. Each participant was tested individually during an approximately 1-hr-long session, and received a cinema ticket voucher in compensation. Age of participants averaged 24.8 years, with a range of 19–42 years.

Materials

A set of 288 Swedish names was constructed. Names were either selected from the set of those at the time (early 2006) popular in the media, or combined (first name plus last name) using frequency tables provided by the national census bureau, Statistics Sweden. The experimenters, when constructing the stimulus material, judged each name as either *famous* or *nonfamous*, and either *frequent* or *infrequent*. Seventy-two names of each of four types were selected. Examples of famous names were: Greta Garbo, Ingmar Bergman, and Björn Borg. The set is available at: www.stenberg.ys.se/Projects/Names/Names.htm.

To verify the judgments, names were checked for *frequency* by looking up each name in the Swedish nationwide telephone directory, and noting the number of hits. This number was log transformed, and used as the variable frequency, which was dichotomized into frequent and infrequent.

Similarly, *celebrity* was checked by making site-specific lookups via the Google search engine. Each name was searched at six Swedish Web sites, affiliated with important media: four national newspapers and two television networks. The number of hits was added across sites, and log transformed. Finally, the variable was dichotomized into famous and nonfamous.

The number of phone directory hits was unrelated to the number of media hits ($r = .01$). The set of names is described in more detail in another publication (Stenberg et al., 2008), where participant ratings corroborating the Internet data are also given.

The set of names was again validated in Experiment 2 of the present study. In this experiment, binary judgments of fame and frequency were made concerning 128 randomly selected names from the set of 288. Although the task was speeded, agreement with the norms was good. The consistency (intraclass correlation) across 13 participants performing the fame judgment task was .96. Proportion correct was .89, and only two items elicited more wrong than correct responses.

The frequency judgment task had to draw a line arbitrarily between what was to be judged common and rare. Our definition was that names with fewer than 10 bearers in Sweden were to be considered rare. The nine participants in this task showed surprising consistency in this rather contrived task; intraclass correlation was .92. Average proportion correct was .84, and eight items, out of 128, were judged incorrectly more than half of the time.

Electrophysiological Recording

The EEG was recorded using tin electrodes in an electrode cap (NeuroScan). Electrodes were placed on the 19 positions of the 10–20 system and referenced to the left mastoid during recording. Additional electrodes were applied to monitor vertical eye movements (VEOG; above and below the left eye), and horizontal eye movements (HEOG; outside the outer canthi). One electrode was applied to the right mastoid and recorded for use in later re-referencing. Electrode sites Fpz and Oz were interpolated from adjacent electrodes.

Amplifiers were set to accept frequencies from 0.1 to 30 Hz, and digitization was performed at a rate of 250 Hz. The data were saved continuously to disk during the session for later off-line processing. The files were visually inspected, and EEG stretches with large artifacts were rejected.

From the continuous EEG files, a template for a typical blink artifact was computed, and corrections for blinks were made to the EEG channels using a regression approach implemented in the NeuroScan software. All EEG channels were re-referenced digitally to an average of the left and the right mastoid. The files were segmented into epochs, consisting of 300 msec prestimulus and 1500 msec poststimulus, and the epochs were baseline-corrected by subtraction of the prestimulus average. An artifact rejection algorithm discarded epochs where any EEG channel deviated from baseline by more than 150 μV. The files were digitally low-pass filtered with a cutoff of 15 Hz and a rolloff of 48 dB. Behavioral data were used to reject trials with incorrect responses. Finally, averages for the different types of names were formed.

Most analyses use a 3 × 3 grid of the following electrodes; F3, Fz, F4, C3, Cz, C4, P3, Pz, P4 from the 10–20 naming convention.

Procedure

The pool of 288 names provided stimuli for four blocks, each with 36 studied names and 36 distractors. The four study–test blocks were run consecutively with subject-terminated pauses between them. Assignment of names to study or distractor status was made randomly for each participant, subject to the constraint that each of the four types of names be represented equally in every block.

Stimuli were presented on a computer screen by an E-prime program. Each name was displayed for 2 sec in the study phase, preceded by a 1-sec fixation cross. In the test phase, a name was first shown for 2 sec while responding was disabled (to give ERPs without motor artifacts). Then, a prompt appeared above the name (''Have you seen this name before in the experiment?''), and two response buttons appeared below it, marked ''Yes'' and ''No.'' A mouse click on either response button terminated this display, which was followed by a new prompt (''How sure are you?'') with three response buttons, marked ''Quite sure,'' ''Relatively sure,'' and ''Not sure.'' A maximum of 7 sec was allowed for response selection, but a response terminated the display, usually much sooner.

Data Processing

Behavioral data. Hit rates and false alarm rates were first computed, using the yes–no responses, and these were used to calculate d' as a measure of accuracy. Further, a 6-point scale, ranging from ''Quite sure new'' to ''Quite sure old'' was constructed from the responses involving confidence and used for the ROC analysis. Yonelinas' model was applied to the ROC data by a Matlab routine written to fit the model to each individual's responses. For each of the four types of name, a value for r, the recollection component, and a value for dp, the signal-detection/familiarity component, were determined. The program can be downloaded from www.stenberg.ys.se/Projects/Names/Names.htm. It works by repeated calls to the standard Matlab function *fminsearch*, and it applies a set of criteria to judge that a satisfactory solution has been reached. Reliability was assessed by testing with varied random starting values, and was found to be .99.

Statistical issues. When testing effects with more than one degree of freedom in the numerator, the Huynh–Feldt correction was applied to the reported probability. The degrees of freedom are given as per before the correction.

Results

Behavioral Results

Analysis of the accuracy data proceeded by first computing the conventional signal detection measures, d' and C, for the four types of names, and submitting them to 2 × 2 ANOVAs, with frequency and celebrity as factors. Sensitivity, d', showed main effects of both frequency [$F(1, 23) = 74.30$, $p < .001$] and celebrity [$F(1, 23) = 83.09$, $p < .001$]. Mean values are shown in Table 1, as are effect sizes. As expected, fame and rarity made names more memorable, and the two factors contributed about equally, with only a marginal interaction between them [$F(1, 23) = 4.19$, $p = .052$].

The criterion, C, was affected only by celebrity [$F(1, 23) = 22.90$, $p < .001$] because a balanced (near zero) criterion was maintained for famous names, whereas the nonfamous were judged against a stricter criterion. There was no main effect of frequency, but an interaction between the two factors [$F(1, 23) = 5.86$, $p = .024$].

The dual-process signal detection model. Application of Yonelinas' DPSD model produced a decomposition of ROC curves which is illustrated in Figure 1. Individual ROCs were fitted and the derived averaged values of r (recollection) and dp (the signal detection component) are reported in Table 1. They were each analyzed in a 2 × 2 ANOVA. On the recollection parameter r, celebrity had a marked effect [$F(1, 23) = 30.98$, $p < .001$, $\eta_p^2 = .57$], producing a much higher probability of recollection for famous names. Frequency had no effect ($F < 1$), nor was there any interaction ($p > .2$).

In contrast, the familiarity parameter, dp, was dominated by a frequency effect [$F(1, 23) = 27.79$, $p < .001$; $\eta_p^2 = .55$], and celebrity had no effect [$F(1, 23) = 2.89$, $p > .10$]. There was no interaction [$F(1, 23) = 1.33$, $p > .20$]. Rare names were more likely to be recognized on the basis of familiarity than were frequent names, as expected by the *familiarity increment* hypothesis.

Table 1. Behavioral Performance in Experiment 1

	Infrequent		*Frequent*		*Effect Size, η_p^2*		
	Nonfamous	*Famous*	*Nonfamous*	*Famous*	*Frequency*	*Celebrity*	*Fr × Ce*
d'	2.00	2.67	1.33	2.28	.76	.78	.15
C	0.25	0.11	0.37	0.06	.03	.50	.20
r	.26	.49	.24	.58	.03	.57	.06
dp	1.61	1.74	.90	1.29	.55	.11	.06

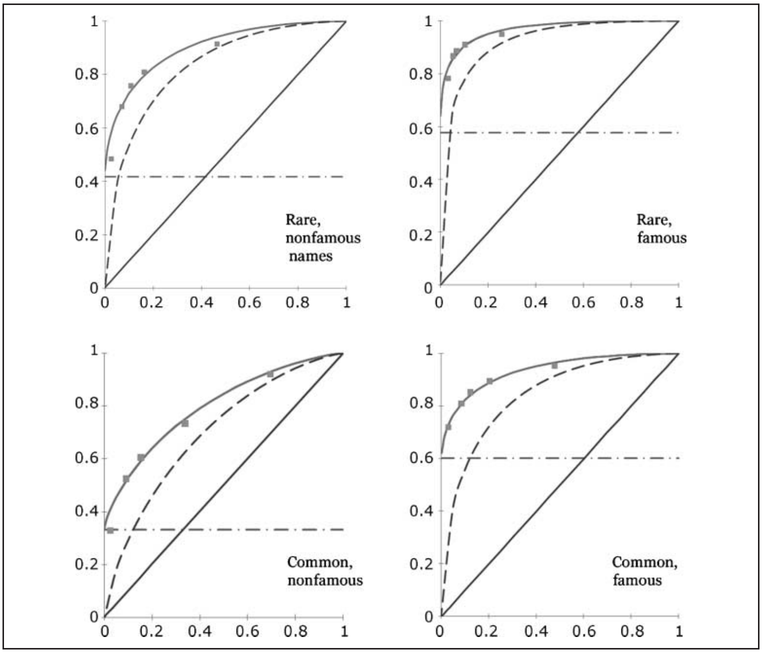


Figure 1. ROC curves for the four types of names. Data points for the summed group ROC are marked by round dots. Solid line indicates model fit. The DPSD model decomposes the ROC into a recollection component, indicated by the horizontal line, and a signal detection component, indicated by the dashed line. Note the relation between fame and recollection, and between frequency and familiarity. (Because this is a group ROC, the values of the parameters do not coincide exactly with the averaged individual values shown in Table 1, but the pattern across conditions is the same.)

Thus, the stimulus qualities frequency and celebrity produced a double dissociation. In terms of Yonelinas's dual-process model, recollection was affected by celebrity alone, and familiarity was affected by frequency alone.

Electrophysiological Results

Our main interest focused on the old–new effects in the test phase. Having seen indicators of familiarity and recollection, derived from Yonelinas' model, align with the stimulus dimensions of our name stimuli, we wondered whether ERP indicators of recollection and familiarity would do so also. We expected the FN400 to covary with frequency and the parietal old–new effect to covary with celebrity.

Old–new effects. In the test phase ERPs, averages were formed for correctly recognized old names, as well as for correctly rejected new ones, within each of the four types defined by frequency and celebrity. The waveforms are displayed in Figures 2 and 3. The average number of trials contributing to each individual average was: F0C0: Hits: 27, Correct Rejections: 32; F0C1: H: 32, CR: 34; F1C0: H: 22, CR: 31; F1C1: H: 31, CR: 31 (where, e.g., F0C1 represents low frequency, high celebrity).

An array of nine electrodes, forming a 3×3 grid over frontal, central, and parietal areas were selected for analysis. The difference waveform was quantified in four intervals: 300–500, 500–700, 700–900, and 900–1100 msec. Interest was directed toward the earliest interval for the frontal effect associated with familiarity, and toward the later ones for the parietal effects associated with recollection. Screening for the absence or presence of an overall old–new effect, we tested the intercept effects first. They were reliable in three of the four intervals [$F(1, 23) = 5.07, p = .034$ in the first; $F(1, 23) = 109.98, p < .001$ in the second; and $F(1, 23) = 18.49, p < .001$ in the third, but $F(1, 23) = 1.21$, *ns* in the last interval].

In the 300–500 msec band, frequency exerted a main effect [$F(1, 23) = 7.76, p = .011, \eta_p^2 = .25$], modified by an interaction with anterior/posterior position [$F(2, 46) = 9.77, p = .001; \eta_p^2 = .30$]. As will be seen in Figures 2 and 4, there were old–new effects for unusual names only, and these were concentrated frontally, with a gradual decline toward the back [$F(1, 23) = 11.85$ for the linear trend contrast]. There were no effects involving celebrity in this time band. In other words, the early frontal effect (FN400) was mainly associated with low frequency names.

In the 500–700 msec interval, there were main effects of both celebrity [$F(1, 23) = 4.70, p = .041, \eta_p^2 = .17$] and, marginally, of frequency [$F(1, 23) = 4.20, p = .052, \eta_p^2 = .16$]. Their spatial gradients were different, although both interacted with the anterior/posterior position

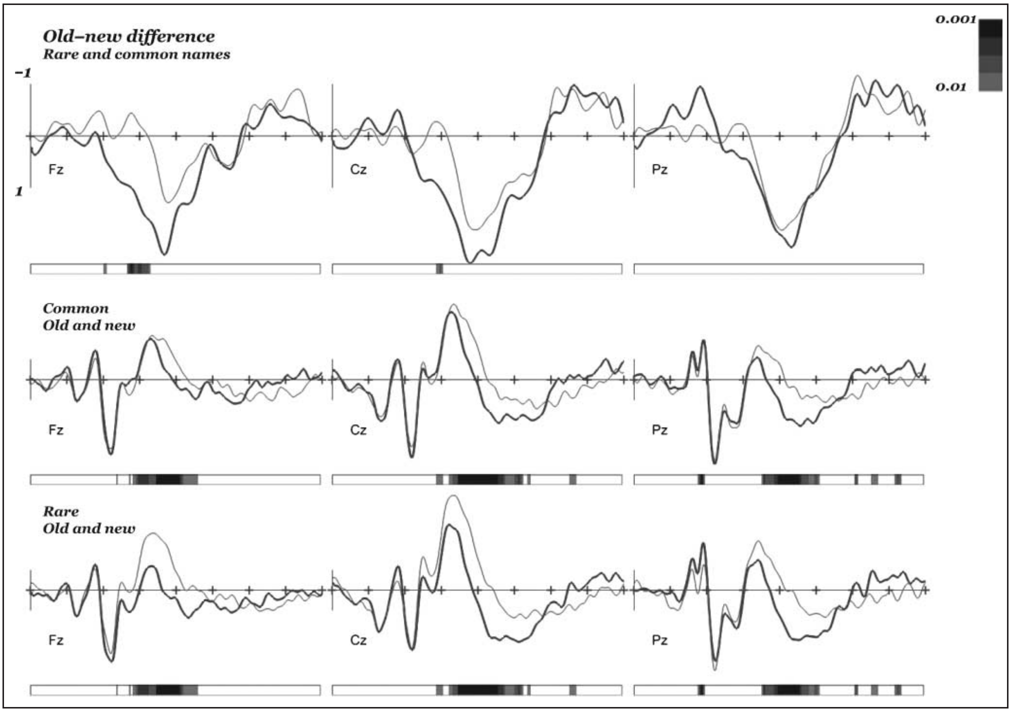


Figure 2. Waveforms in the test phase. Bottom row: Rare names. Thin line: new, correctly rejected names. Bold line: old, correctly endorsed names. Second row: Frequent names, new and old. Top row: Difference waves: old minus new. Bold line: rare names; thin line: frequent names. Electrode locations Fz (frontal), Cz (central), and Pz (parietal) on the midline. In each panel, the bar at the bottom shows the outcome of a sequence of *t* tests, one every 20 msec, of the difference between the two waveforms. Dark bars indicate probabilities <.01. Epoch −200 to 1400 msec, a marker each 200 msec. Vertical line at the beginning of each trace shows −1 μV (upward) and +1 μV (downward).

[$F(2, 46) = 5.12$, $p = .020$, $\eta_p^2 = .18$ for frequency and $F(2, 46) = 11.00$, $p = .001$, $\eta_p^2 = .32$ for celebrity]. This proved to be due to opposite linear trends because frequency had its largest effect frontally, $\eta_p^2 = .27$ and a gradual wearing-off toward the back: $\eta_p^2 = .16$ and $\eta_p^2 = .02$ for the central and parietal electrodes, respectively. In contrast, the celebrity effect grew stronger at posterior sites: $\eta_p^2 = .01$, $\eta_p^2 = .20$, and $\eta_p^2 = .33$, for the frontal, central, and parietal sites, respectively. This is illustrated in Figures 2, 3, and 4. The parietal old–new effect was thus most strongly associated with high celebrity names (Figure 5).

In the 700–900 msec interval, no main effects of frequency and celebrity were reliable, and only marginally reliable interactions were found between celebrity and the electrode position factors, $.05 < p < .1$, which will not be pursued further. In the 900–1100 msec range, however, the slow posterior wave seen in other high-demand memory tasks arose (Johansson & Mecklinger, 2003). There was a main effect of celebrity [$F(1, 23) = 8.82$, $p = .007$], as well as a marginal interaction with anterior/posterior position [$F(2, 46) = 3.53$, $p = .059$], and an interaction, Frequency × Celebrity [$F(1, 23) = 7.17$, $p = .013$]. A follow-up test of the parietal row of electrodes showed a main effect of celebrity [$F(1, 23) = 14.28$, $p = .001$], and again an interaction with frequency [$F(1, 23) = 4.66$, $p = .041$], indicating that a late posterior effect accompanied both famous and common names, that is, all types except the infrequent, unheard-of names, for which source attribution was not an issue.

Confidence. To assess the effect of confidence on the FN400 and the parietal positive component, trials were sorted into hits and correct rejections, and within each of those categories into high and low confidence responses. Of special interest were the correct rejections. Both theory and earlier investigations (Woodruff et al., 2006) suggest that familiarity is a graded process, correlated with confidence. Recollection is expected to be a thresholded phenomenon, hence, impervious to confidence effects on negative decisions (Curran, 2004). For hits, pairwise *t* tests showed effects of confidence on both the

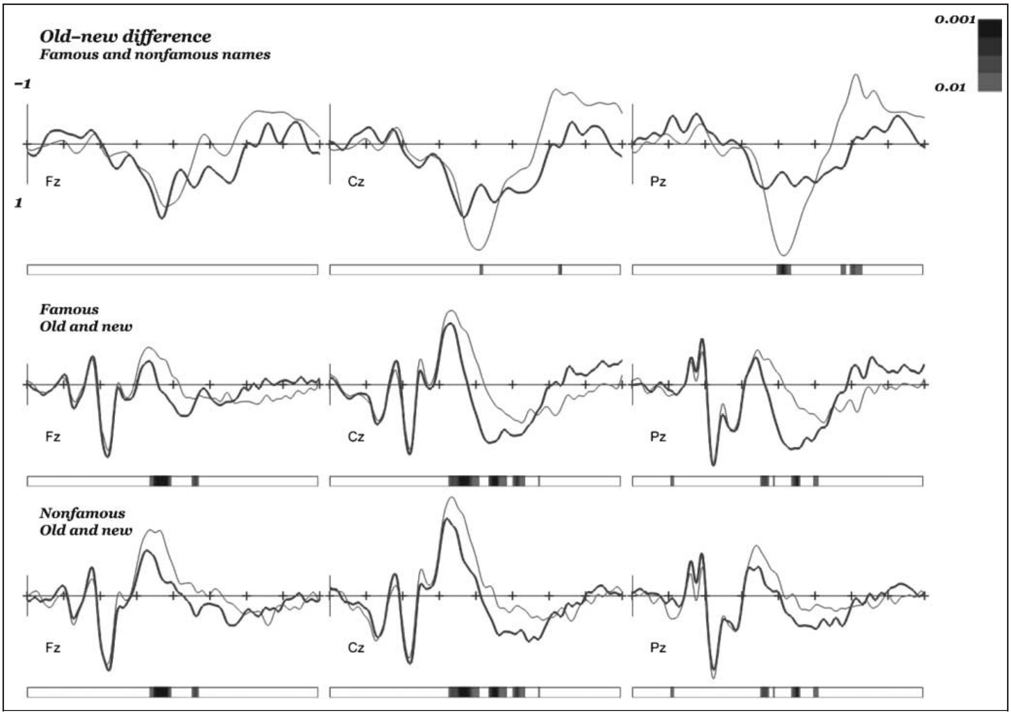


Figure 3. Waveforms in the test phase. Nonfamous (bottom) versus famous (middle row) names. Otherwise as in Figure 2.

FN400 (at Fz, 300–500 msec) and the parietal positive component (at P3, 500–700 msec) [$t(22) = 2.14$, $p = .043$, and $t(22) = 3.87$, $p = .001$, respectively]. More importantly, correct rejections differed according to confidence only for the FN400 [$t(23) = 2.10$, $p = .047$], and not for the parietal positive component [$t(23) < 1$, *ns*; see Figure 6].

Discussion

So far, we have seen how the dimensions of celebrity and familiarity could separate behavioral responses as well as test-phase ERPs along the dividing lines delineated by the two-process theory of recognition. Still, the possibility remains that we have confounded priming with familiarity. We therefore performed a second experiment, in which our stimulus material was used in a task where priming could be measured. The purpose was to see if priming covaried with one of our stimulus dimensions, and if so, which one.

Our expectations are that famous names produce more priming than nonfamous ones. Because they have a clear semantic content, a host of associations can be activated by them, helping to produce facilitation in a renewed encounter. Nonfamous names, on the other hand, evoke few images, especially if they are rare. A faceless name has no semantics and should produce little conceptual priming.

EXPERIMENT 2

We needed two tasks that could be meaningfully applied to the whole of our stimulus material and could be trusted to activate the conceptual content of the names. Our choice fell upon a frequency judgment task (''Is this a common name?'') and a celebrity judgment task (''Is this the name of a famous person?''). Two different tasks, one for the study phase and one for the test phase, were used to avoid measuring just the learning of a fixed stimulus–response mapping.

Responding was speeded, with reaction time as the dependent variable. For one group of participants, the frequency judgment was presented first, in the incidental study phase, and the celebrity task thereafter, at test, when the degree of priming was measured. This will be called the FC condition. For another group (the CF condition), the order was reversed, and priming was measured in the final frequency judgment task. The degree of priming was measured in the second task as the

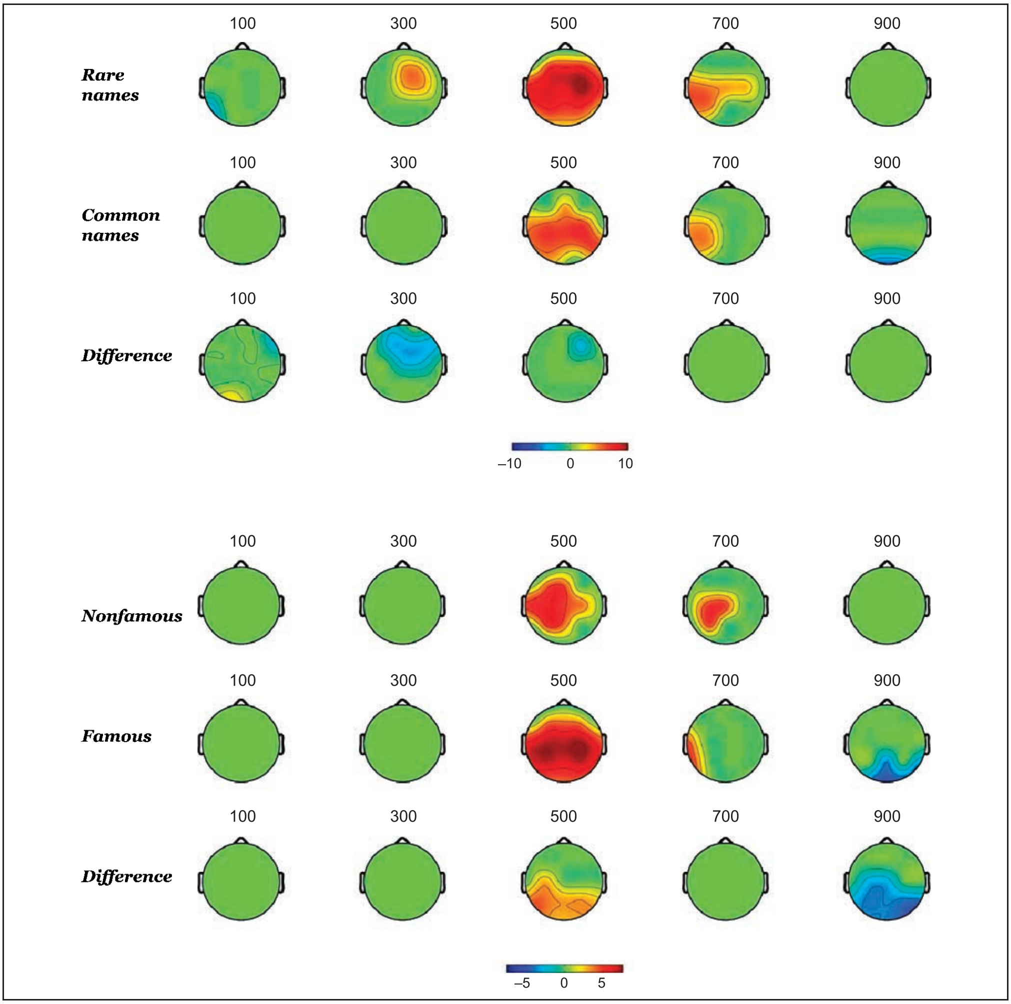


Figure 4. Topographical maps (schematic head seen from above, front upward) showing old–new effects. First and second rows show t values testing whether the old–new effect is different from zero. t Values under a threshold corresponding to $p = .001$ have been set equal to zero. Third row shows t tests of the difference between high and low frequency names, using a threshold of $p = .01$. Rows four to six show similar t tests for high and low celebrity names. Dark areas with bright contours correspond to positive t values, bright areas with dark contour lines correspond to negative values.

difference in reaction time between re-presented names and new names.

Methods

The experiment was performed via the Internet, using the Inquisit software (www.millisecond.com). Twenty-two participants were recruited from a pool of volunteer students (two-thirds women, mean age = 25 years) and university faculty. A set of 64 names was presented in the first task, and the same 64 plus 64 new names were presented in the second task. The assignment of names to either group was counterbalanced across subjects. The whole set of 128 names was randomly selected from our pool of 288, with the restriction that equal numbers of the four types of names be present. Outlier reaction times (with a cutoff at the individual mean ± 2.5 *SD*s) were replaced by the cutoff value, and priming scores were

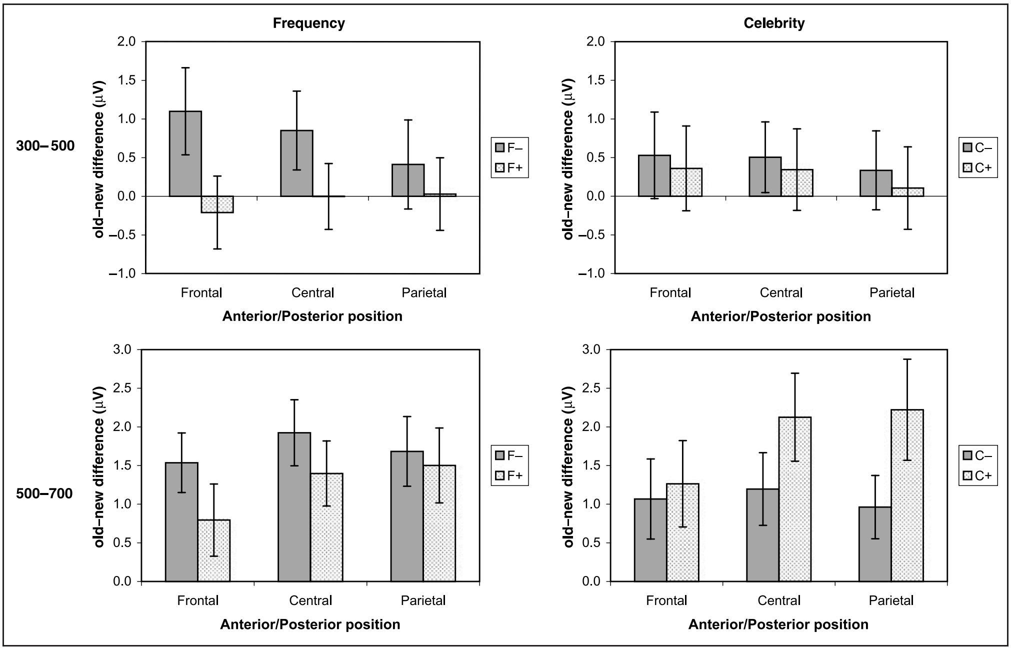


Figure 5. Means and 95% confidence intervals for the old–new effect in the 300–500 and 500–700 msec intervals. Left panels: high and low frequency names. Right panels: high and low celebrity names. Note the old–new effect in the early interval frontally for rare names (F−). In the later interval there are old–new effects for all types of names, but a maximum can be found parietally for famous names (C+).

formed for the four types of names by subtraction of old-name reaction times from new-name reaction times.

Results and Discussion

Priming scores (new-name RTs minus old-name RTs) were subjected to a 2 × 2 × 2 (Frequency × Celebrity × Task Order) analysis. The main effect of celebrity was significant [$F(1, 20) = 8.13$, $p = .01$] because famous names evinced priming ($M = 36$ msec, 95% CI = 12–60) and nonfamous ones did not. The effect of frequency, and the interaction, was nonsignificant (both $F < 1$). There was no main effect of task order, but it modified the effect of celebrity through an interaction [$F(1, 20) = 5.36$, $p = .031$]. The simple main effect of celebrity was significant in the FC task order [$F(1, 12) = 26.60$, $p < .001$]. In other words, if celebrity was salient when priming was measured, it evinced a large priming effect (61 msec; CI = 23–99, for famous names). On the other hand, when frequency was made salient at test, that is, in the CF task order, neither frequency nor celebrity promoted any priming: all Fs < 1 (priming scores for the two task orders are shown in Table 2).

In Figure 7, priming scores (averaged across task orders) are shown along with FN400 amplitudes (mean of Fz in 300–500 msec). Conceptual priming was sensitive only to celebrity, and the FN400 only to frequency, that is, they were completely independent.

The present two tasks (speeded frequency and celebrity decisions) used to assess conceptual priming were of the passive verification type. Typical conceptual priming tasks, in contrast, are often production tasks, such as category exemplar generation, and it has been shown that production and verification tasks can yield different results because of different degrees of response competition (Vaidya & Gabrieli, 2000). However, the recognition task in which the FN400 has been found is also of the verification type. Further, the same perceptual form was repeated in the study and test phases of our conceptual priming experiment, and therefore, it cannot be excluded that perceptual priming also contributed. Against that possibility speaks the fact that there was no priming in the CF task order, and in any case, our tasks share the perceptual repetition factor with most recognition memory experiments.

GENERAL DISCUSSION

In summary, we found that our stimulus dimensions dissociated familiarity and recollection. Recollection was

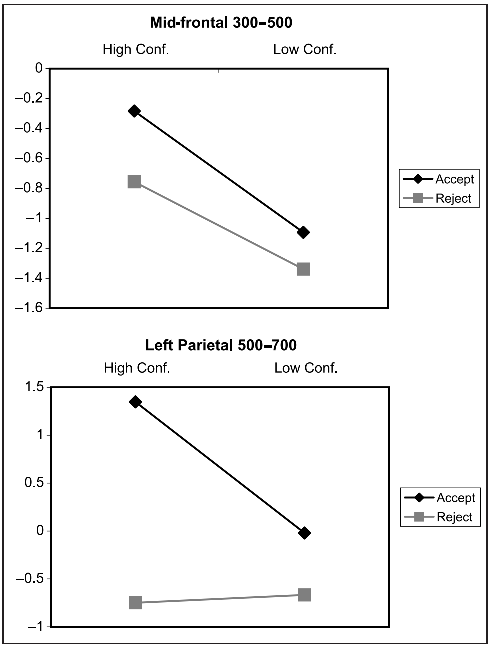


Figure 6. The FN400 and the parietal positive component as a function of confidence and decision (Accept, Reject) in hits and correct rejection. Note that in rejections, FN400 amplitude covaries with confidence, in contrast to the parietal component.

accompanied by the expected parietal old–new effect, as well as a later, posterior negativity, often found in source memory tasks. Familiarity was accompanied by the mid-frontal FN400 effect, which, importantly, was orthogonal to conceptual priming.

The behavioral results of this study confirmed the expectation that famous names would be better remem-

Table 2. Priming Scores in Experiment 2

Task Order	*Type of Name*	*M*	*SE*
CF (frequency task at test)	Rare, nonfamous	7.7	39.0
	Rare, famous	−0.8	21.9
	Frequent, nonfamous	−4.2	18.7
	Frequent, famous	21.4	20.7
FC (celebrity task at test)	Rare, nonfamous	−12.6	32.5
	Rare, famous	78.8[a]	18.2
	Frequent, nonfamous	−31.1	15.6
	Frequent, famous	43.2[a]	17.2

[a]Indicates significant priming.

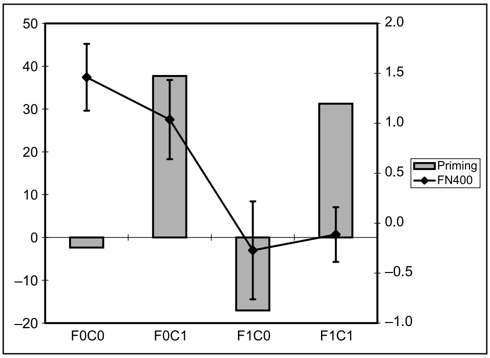


Figure 7. FN400 amplitudes (lines) and conceptual priming scores (bars) across the four types of names, frequency (F) and celebrity (C) to a low (0) or high (1) degree. Error bars equal 1 *SEM*.

bered than nonfamous ones, and that infrequent names would be better remembered than frequent ones. Crucially, the degree of recollection covaried with the celebrity of the names, and independently, the degree of familiarity covaried with the frequency. In stating this, we rely on analysis according to Yonelinas's DPSD model, which provided parameter values for each name type and each individual. The model rests on the assumption of a high-threshold process underlying recollection, and an equal-variance signal detection process underlying familiarity. Both of these assumptions have been questioned, especially the first. Alternative proposals are afoot, such as single-process, unequal-variance models (Heathcote, Raymond, & Dunn, 2006), and dual-process models which treat both as continuous signal detection processes (Wixted, 2007; Rotello, Macmillan, & Reeder, 2004). The viability of the DPSD and its competitors is, at present, a topic under intense investigation (Yonelinas & Parks, 2007). Our mapping between the name dimensions and memory processes does not rest on the DPSD model alone. In our earlier article (Stenberg et al., 2008), we have given several reasons for this mapping, none of which depended upon the assumptions of the DPSD model. We noted that the celebrity dimension affected primarily hit rates, and frequency affected false alarm rates. This pattern is indicative of the workings of two processes (Reder et al., 2000). Moreover, the slopes of the ROCs were affected by celebrity, not by frequency. In toto, the data indicate that the dimensions of the name stimuli align with familiarity and recollection.

The parietal old–new effect proved sensitive to the celebrity of the names. With increasing fame, the network of known facts and associations grows larger, and with this expanded net, the catch increases. The amount of retrieved information grows with contributions from both semantic and episodic memory. It is likely that the amplitude of the parietal effect reflects this abundance (Wilding, 2000). Faced with the task of deciding if the retrieved material derives from the experiment proper or from the media, he engages in attempts to reconstruct the context in which the name was recently encountered. This contextual binding effort forms one of the two conditions in which, according to a recent review (Johansson & Mecklinger, 2003), the late posterior negativity follows the positive parietal old–new effect (see Friedman, Cycowicz, & Bersick, 2005 for a related view). In our data, famous names evoked the largest positive (500–700) old–new effect, and they also produced the largest negative (900–1100) amplitude modulation. This is consistent with a mechanism that retrieves a mixture of memory facts and episodes, and if necessary, engages in effortful reconstruction of the links from these to their context. With our task, mere quantity of retrieved information was not enough. A definite bond to the context of the experiment was needed to make the information useful. Similarly, task demands for retrieval of special circumstances at encoding, or a ''remember'' decision have been seen to produce the late, posterior negativity (Mecklinger, Johansson, Parra, & Hanslmayr, 2007; Wolk et al., 2007).

The mid-frontal effect, the FN400, increased for rare names relative to common names, as did recognition by familiarity. In line with expectations and previous findings, the FN400 proved sensitive in a graded fashion to the strength or confidence (Woodruff et al., 2006; Finnigan, Humphreys, Dennis, & Geffen, 2002) of both positive and negative decisions, in contrast to the parietal positive component (Curran, 2004), which did not show a graded response in ''reject'' decisions. The latter finding is predictable from the theoretical position that there can be no below-threshold recollection. Familiarity, on the other hand, can take on any value on a continuous scale.

Our findings are entirely compatible with the view that the FN400 is a reflection of the processes underlying familiarity. The alternative hypothesis, connecting the FN400 with conceptual priming, is difficult to reconcile with the fact that rare names evoked a larger FN400 than common names, and that celebrity made no difference. It would seem that rare names give little opportunity for conceptual processing. Especially the kind of rare names not belonging to a famous person would seem to be the equivalent of pseudowords and nonfigurative doodles, that is, stimuli without meaning. Where there is no conceptual content, there can be no conceptual priming.

Moreover, our data from Experiment 2 are quite unequivocal in showing that priming was sustained by fame, and fame alone. When fame was made salient by the speeded judgment task, famous names produced a large priming effect. When frequency was made salient in the same way, there were no priming effects for any type of name.

Coupled with the completely different relation for the FN400, which was a function of frequency, and frequency alone, these data point unambiguously in one direction. Two independent processes contribute to recognition: familiarity, which is reflected in the FN400, and recollection, which is reflected in two later, parietal components.

Acknowledgments

This research was supported by grants from the Swedish Research Council to Georg Stenberg (''Det semantiska minnet''), Mikael Johansson (VR 60229601), and Ingmar Rosén (VR 0084).

Reprint requests should be sent to Georg Stenberg, Department of Psychology, School of Behavioural Sciences, Kristianstad University, SE-291 88 Kristianstad, Sweden, or via e-mail: georg.stenberg@hkr.se.

REFERENCES

Aggleton, J. P., & Brown, M. W. (2006). Interleaving brain systems for episodic and recognition memory. *Trends in Cognitive Sciences, 10,* 455–463.

Brown, M. W., & Aggleton, J. P. (2001). Recognition memory: What are the roles of the perirhinal cortex and hippocampus? *Nature Reviews Neuroscience, 2,* 51–61.

Clark, S. E., & Gronlund, S. D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin & Review, 3,* 37–60.

Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition, 28,* 923–938.

Curran, T. (2004). Effects of attention and confidence on the hypothesized ERP correlates of recollection and familiarity. *Neuropsychologia, 42,* 1088–1106.

Curran, T., & Cleary, A. M. (2003). Using ERPs to dissociate recollection from familiarity in picture recognition. *Cognitive Brain Research, 15,* 191–205.

Curran, T., DeBuse, C., & Leynes, P. A. (2007). Conflict and criterion setting in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 2–17.

Cycowicz, Y. M., Friedman, D., & Snodgrass, J. G. (2001). Remembering the color of objects: An ERP investigation of source memory. *Cerebral Cortex, 11,* 322–334.

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,* 379–386.

Duzel, E., Vargha-Khadem, F., Heinze, H. J., & Mishkin, M. (2001). Brain activity evidence for recognition without recollection after early hippocampal damage. *Proceedings of the National Academy of Sciences, U.S.A., 98,* 8101–8106.

Fernandez, G., & Tendolkar, I. (2006). The rhinal cortex: ''gatekeeper'' of the declarative memory system. *Trends in Cognitive Sciences, 10,* 358–362.

Finnigan, S., Humphreys, M. S., Dennis, S., & Geffen, G. (2002). ERP ''old/new'' effects: Memory strength and decisional factor(s). *Neuropsychologia, 40,* 2288–2304.

Friedman, D., Cycowicz, Y. M., & Bersick, M. (2005). The late negative episodic memory effect: The effect of recapitulating study details at test. *Brain Research, Cognitive Brain Research, 23,* 185–198.

Gardiner, J. M., & Richardson-Klavehn, A. (2000). Remembering and knowing. In E. Tulving & F. Craik (Eds.), *The Oxford handbook of memory* (pp. 229–244). Oxford: Oxford University Press.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91,* 1–67.

Heathcote, A., Raymond, F., & Dunn, J. (2006). Recollection and familiarity in recognition memory: Evidence from ROC curves. *Journal of Memory and Language, 55,* 495–514.

Herron, J. E. (2007). Decomposition of the ERP late posterior negativity: Effects of retrieval and response fluency. *Psychophysiology, 44,* 233–244.
Hintzman, D. L. (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments & Computers, 16,* 96–101.
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,* 1–18.
Johansson, M., & Mecklinger, A. (2003). The late posterior negativity in ERP studies of episodic memory: Action monitoring and retrieval of attribute conjunctions. *Biological Psychology, 64,* 91–117.
Johansson, M., Stenberg, G., Lindgren, M., & Rosen, I. (2002). Memory for perceived and imagined pictures—An event-related potential study. *Neuropsychologia, 40,* 986–1002.
Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87,* 252–271.
Mecklinger, A. (2006). Electrophysiological measures of familiarity memory. *Clinical EEG and Neuroscience, 37,* 292–299.
Mecklinger, A., Johansson, M., Parra, M., & Hanslmayr, S. (2007). Source-retrieval requirements influence late ERP and EEG memory effects. *Brain Research, 1172,* 110–123.
Murdock, B. B. (1993). TODAM2: A model for the storage and retrieval of item, associative, and serial-order information. *Psychological Review, 100,* 183–203.
Nessler, D., & Mecklinger, A. (2003). ERP correlates of true and false recognition after different retention delays: Stimulus- and response-related processes. *Psychophysiology, 40,* 146–159.
O'Reilly, R. C., & Norman, K. A. (2002). Hippocampal and neocortical contributions to memory: Advances in the complementary learning systems framework. *Trends in Cognitive Sciences, 6,* 505–510.
Paller, K. A., Kutas, M., & McIsaac, H. K. (1995). Monitoring conscious recollection via the electrical activity of the brain. *Psychological Science, 6,* 107–111.
Paller, K. A., Voss, J. L., & Boehm, S. G. (2007). Validating neural correlates of familiarity. *Trends in Cognitive Sciences, 11,* 243–250.
Parks, C. M., & Yonelinas, A. P. (2007). Moving beyond pure signal-detection models: Comment on Wixted (2007). *Psychological Review, 114,* 188–202; discussion 189–203.
Reder, L. M., Nhouyvanisvong, A., Schunn, C. D., Ayers, M. S., Angstadt, P., & Hiraki, K. (2000). A mechanistic account of the mirror effect for word frequency: A computational model of remember–know judgments in a continuous recognition paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 294–320.
Rotello, C. M., Macmillan, N. A., & Reeder, J. A. (2004). Sum–difference theory of remembering and knowing: A two-dimensional signal-detection model. *Psychological Review, 111,* 588–616.
Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences, 11,* 251–257.
Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A. M., Birch, C. S., & Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature, 392,* 595–598.
Rugg, M. D., & Yonelinas, A. P. (2003). Human recognition memory: A cognitive neuroscience perspective. *Trends in Cognitive Sciences, 7,* 313–319.
Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM-retrieving effectively from memory. *Psychonomic Bulletin & Review, 4,* 145–166.
Skinner, E. I., & Fernandes, M. A. (2007). Neural correlates of recollection and familiarity: A review of neuroimaging and patient data. *Neuropsychologia, 45,* 2163–2179.
Stenberg, G., Hellman, J., & Johansson, M. (2008). The memorability of names and the divergent effects of prior experience. *European Journal of Cognitive Psychology, 20,* 312–345.
Vaidya, C. J., & Gabrieli, J. D. E. (2000). Picture superiority in conceptual memory: Dissociative effects of encoding and retrieval tasks. *Memory & Cognition, 28,* 1165–1172.
Voss, J. L., & Paller, K. A. (2006). Fluent conceptual processing and explicit memory for faces are electrophysiologically distinct. *Journal of Neuroscience, 26,* 926–933.
Wilding, E. L. (2000). In what way does the parietal ERP old/new effect index recollection? *International Journal of Psychophysiology, 35,* 81–87.
Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review, 114,* 152–176.
Wolk, D. A., Schacter, D. L., Lygizos, M., Sen, N. M., Chong, H., Holcomb, P. J., et al. (2007). ERP correlates of Remember/Know decisions: Association with the late posterior negativity. *Biological Psychology, 75,* 131–135.
Woodruff, C. C., Hayama, H. R., & Rugg, M. D. (2006). Electrophysiological dissociation of the neural correlates of recollection and familiarity. *Brain Research, 1100,* 125–135.
Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1341–1354.
Yonelinas, A. P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. *Memory & Cognition, 25,* 747–763.
Yonelinas, A. P. (1999). The contribution of recollection and familiarity to recognition and source-memory judgments: A formal dual-process model and an analysis of receiver operating characteristics. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 1415–1434.
Yonelinas, A. P. (2001a). Components of episodic memory: The contribution of recollection and familiarity. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 356,* 1363–1374.
Yonelinas, A. P. (2001b). Consciousness, control, and confidence: The 3 Cs of recognition memory. *Journal of Experimental Psychology: General, 130,* 361–379.
Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46,* 441–517.
Yonelinas, A. P., Kroll, N. E. A., Dobbins, I., Lazzara, M., & Knight, R. T. (1998). Recollection and familiarity deficits in amnesia: Convergence of remember–know, process dissociation, and receiver operating characteristic data. *Neuropsychology, 12,* 323–339.
Yonelinas, A. P., & Parks, C. M. (2007). Receiver operating characteristics (ROCs) in recognition memory. *Psychological Bulletin, 133,* 800–832.
Yonelinas, A. P., Quamme, J. R., Widaman, K. F., Kroll, N. E., Sauve, M. J., & Knight, R. T. (2004). Mild hypoxia disrupts recollection, not familiarity. *Cognitive, Affective & Behavioral Neuroscience, 4,* 393–400; discussion 401–406.
Yonelinas, A. P., Widaman, K., Mungas, D., Reed, B., Weiner, M. W., & Chui, H. C. (2007). Memory in the aging brain: Doubly dissociating the contribution of the hippocampus and entorhinal cortex. *Hippocampus, 17,* 1134–1140.
Yovel, G., & Paller, K. A. (2004). The neural basis of the butcher-on-the-bus phenomenon: When a face seems familiar but is not remembered. *Neuroimage, 21,* 789–800.

Paper III

The Variance Theory of Recognition Memory

Johan Hellman[1], Sverker Sikström[1]

[1]Lund University, Sweden

In current popular models of recognition memory, the influence of pre-experimental familiarity is rarely described in detail. In the current paper, we used the variance theory (Sikström, 2001), which is directly based on the effect of pre-experimental familiarity. The model is implemented in a two-layer network, representing items and contexts, respectively. The pre-experimental encoding influences variability of the input to respective layer, the number of active nodes, and thereby, recognition strength. It is shown that the non-linear activation function is an essential mechanism for accounting for the relevant psychological phenomena. The model was applied to recognition memory data collected with the name paradigm (Stenberg, Hellman & Johansson, 2008) with fame and frequency as independent variables, in source memory and paired-associates tasks. Frequency increased the variability of the net input to the context layer, yielding better performance for low than high frequent names, but with no effect on source memory or associative memory. Fame increased the percentage of active nodes on both layers, but not the variability, resulting in higher performance in item memory and a reliable effect on both source and associative memory. The variance theory is discussed in the context of signal detection theory and dual process theory.

Acknowledgements

This research was supported by the Swedish Research Council (grant to project "Semantic Spaces as a Method in Psychology").

During several decades, the study of recognition memory has gained an important place in the memory literature (Yonelinas, 2002; Achilles 1920; Clark & Gronlund, 1996) because the investigation of a binary memory decision reveals a multitude of insights of memory, and because recognition memory is pervious to the methodology used to study the cognitive neuroscience of memory. This interest has resulted in several more or less comprehensive models of recognition memory. These models have evolved from signal detection theory (Egan, 1958) to several global matching models and single- and dual process models of recognition memory (Clark & Gronlund, 1996; Yonelinas & Parks, 2007). In the present paper, a model that have provided a feasible account of the mirror effect (Sikström, 2001) will be generalized and implemented on item, source and associative memory using a paradigm developed to study the contribution of prior knowledge on episodic memory, namely the name paradigm (Stenberg, Hellman & Johansson, 2008; Stenberg et al., 2009). The variance theory (denoted the VT) directly models how pre-experimental frequency and fame influences variability in item and context layers and is therefore of particular interest for understanding frequency and familiarity effects in recognition performance. Further, effects of fame on recognition memory are poorly investigated (Bowles et al., 2012). We therefore focus our theoretical understanding of the data in this article on this model. The VT is tested in two experiments based on the name paradigm, using the source memory framework and the paired associates test. By using the VT to account for data recorded with the name paradigm, we aim to provide a more detailed account of fame and frequency effects, and how pre-experimental familiarity affects recognition memory.

Prior knowledge and recognition memory

The effect of prior knowledge on memory has been shown to influence recognition memory decisions for words (Estes & Maddox, 2002; Reder et al., 2000), faces (Bird & Burgess, 2008; Bird, Davies, Ward, & Burgess, 2011) and names (Stenberg, Hellman, & Johansson, 2008; Stenberg et al., 2009). Further, when comparing experts and novices in memory for modality specific knowledge, memory performance differs reliably for position of chess pieces (de Groot, Gobet & Jongman, 1996), wine related odors (Parr, Heatherbell, & White, 2002) and odor compounds associated with beer beverages (Valentin, 2007). In a comparison of memory for personally known and unknown faces, it was shown that personally known faces were associated with an increase in the separation of the old and new item distributions in comparison with personally unknown faces – resulting in higher memory performance in terms of hits and correct rejections (Bird et al., 2011). Further, Stenberg and colleagues investigated memory for famous and non-famous names with high and low frequency in the recently developed name paradigm. Orthogonal variations of the fame and frequency variables were used as test material in recognition memory experiments. Whereas fame increased memory strength (i.e., increasing d'), frequency had a detrimental effect, which is interesting because both fame and frequency reflects prior experience and therefore should be expected to elevate memory performance (Anderson & Schooler, 1991; Dennis, 1995). On the other hand, it is a ubiquitous finding that high frequency impairs recognition memory in comparison with memory for rare items (the word frequency effect), where the former increases false alarms and lowers hits, whereas the latter has the opposite effect, a pattern known as the mirror-effect (Glanzer & Adams, 1990a). Thus, the observation that prior knowledge should improve memory and the empirical finding of the word-frequency effect is contradictory.

The Name paradigm

The name paradigm emanates from the idea that pre-experimental knowledge, such as the knowledge that a name is frequent or infrequent and that a person is famous or non-famous, affects recognition in different ways. Thus, the participant is presented with and instructed to remember famous names, both frequent (e.g., using English equivalents, Tom Jones, Gordon Brown) and infrequent (e.g. Gwyneth Paltrow, Javier Bardem), as well as non-famous names, also frequent (e.g. John Smith, Jane Cooper) and infrequent (e.g. Sebastian Weisdorf, Brogan Kincaid). Stenberg and colleagues demonstrated that the two stimulus dimensions not only affect accuracy, inducing a mirror effect in both studies (Stenberg, Hellman & Johansson, 2008; Stenberg et al., 2009). The authors also reported a double dissociation of hits and false alarms, and that the two variables were related to different electrophysiological correlates of recognition (Stenberg et al., 2009).

Stenberg and colleagues proposed that fame affects recognition differently than frequency because a famous name is endowed with a richness of associations; the name Gwyneth Paltrow is associated with knowledge about a face, appearances in different movies and so forth. This activated semantic knowledge enhances encoding by providing links to individuating features of the name bearer, and lead to better discriminability at test. These conceptual features lead not only to a facilitation in differentiation of new and old items at test, but also makes the names less susceptible to interference and/or provide more features that can be cued in a recognition test.

Frequency, on the other hand, had a detrimental effect on recognition performance. Because a frequent name is unrelated to a specific face, voice or any other individuating feature at test, it becomes difficult to determine whether the item was presented during the study phase of the experiment, or in any other event prior to the experiment. Further, infrequent names were better recognized than the frequent ones.

The odd character of the infrequent names result in focus of attention, and the subsequent increment in memory strength increases the probability for an old response in the test phase (Mandler, 1980; 2008). Because the participant most likely has no prior experience of such a name bearer, there is no pre-experimental contextual information to relate the name to.

Previous findings with the name paradigm have been accounted for by relating fame and frequency to different memory processes (familiarity and recollection) in accordance with dual-process theory (Yonelinas & Parks, 2007). In the current article, we want to account for fame and frequency effects with the variance theory (Sikström, 2001), and thereby present a more detailed computational account of the paradigm.

The Variance Theory

The variance theory has successfully accounted for the mirror effect of recognition memory (Sikström, 2001). A short description of the model follows below, and a mathematical specification is given in the Appendix (for a detailed description of the model, see Sikström, 2001).

Representation

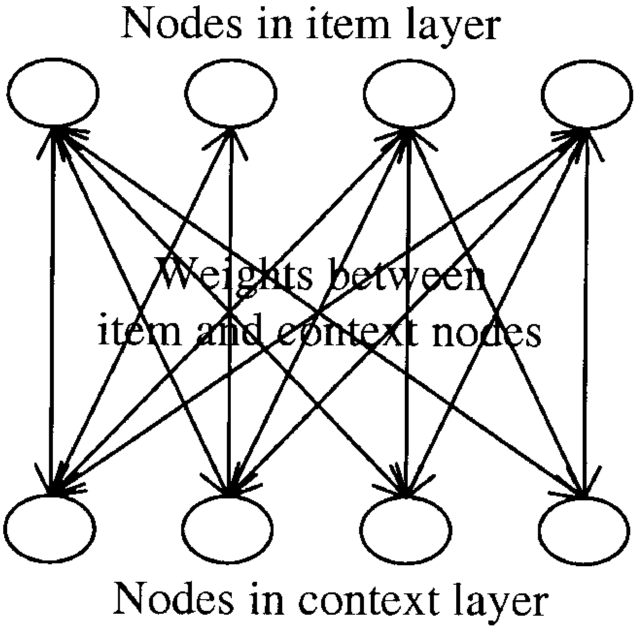
The model constitutes a neural network that encompasses two layers in the memory representations, an item layer and a context layer (see Figure 1), and each layer influences the other layer. Vectors of binary features represent items and contexts, where a node represents each feature. The two layers are fully inter-connected, with weights attached to each connection between nodes in the item and context layers; however, there are no connections within the layers. Respective layer receives input from the other one at recognition, where the activation patterns of the features during study are reinstated. Thus, input to the context layer affects the item layer, and input to the item layer affects the context layer.

Encoding

During encoding an item is presented in a certain encoding context and the weight between item and context features are modified according to a Hebbian learning rule, which increase the connection weights between simultaneously activated nodes. When an item is encoded, it is associated with pre-experimental contexts, and the number of contexts it is associated with is determined by level of item frequency.

Thus, a low frequent item is associated with a smaller number of contexts as compared to a high frequent item, which influence how much pre-experimental encoding is confused with experimental encoding.

Figure 1. Architecture of the Variance Theory



Recognition

An old response is made if the number of active features exceeds a threshold. At retrieval, the to-be-recognized item and the context are reinstated as cues. The feature at the context layer serves as cues for the item features, and the item features serves as cues for the context features. A feature is activated if the net input to the feature exceeds a threshold, and is active in the cued representation (i.e., information about the target and context that is presented at retrieval). The net input is calculated by summing all weights from active features, thus the net input of an item feature is the sum of the weights connected to active context features, and the net input to a context feature is the sum of the weights connected to active item features.

For new items, the net input equals the sum of random weights, where the expected values of those weights are zero. For old items, the expected net input will have an expected value above zero if both nodes were active at encoding, due to the learning rule applied at encoding (i.e., if the connection between nodes has been strengthen during encoding).

Theoretical aspects of variability in the VT

In the VT, familiarity depends on the non-linear activation function, which either activates a feature or not. The non-linearity introduces a difference in the probability for old and new features to be active, where old items have a larger probability of being active than new items, and therefore show a larger variability in familiarity. The probability density function of active nodes can be described by a binomial distribution, where old items have larger variability than new items, because the former is associated with a higher probability than the latter.

The probability that the binary features are activated increase with the strength of the input but also depends on the variability of the input. The presented test item produces a net input to the context layer, which encompass input from both the study context and from several pre-experimental contexts associated with the item. Thus, information about items experienced in the experimental study context and previous events is kept in the context layer. The variability in net input to the context nodes from the item nodes increase linearly with the number of contexts associated with the item, and vice versa, the variability in net input to the item nodes increase with the number of items associated with the context. The probability that the features are activated increase with the variability of the input for new items and decreases for old items. At recognition, the presented test item leads to a re - instantiation of encoded information from previously experienced events. Thus, variability in the network depends on how many contexts are associated with the test item and the number of contexts associated with an item.

Frequency and fame

Recognition performance of high and low fame and frequency can be related to activity in the item and context layers. High frequency items are associated with more pre-experimental contexts, resulting in larger variability in the context representation of the item, which leads to lower familiarity for old high than old low frequency items (due to the increased variability in net input). Because high frequent names are associated with more contexts, it is difficult for the subject to differentiate the pre-experimental familiarity for the name, and the familiarity induced by the study context. The variability of the net inputs in the context nodes increases linearly with the number of times an item has been pre-experimentally studied in different contexts. Thus, the variability in net input to context nodes for high frequent items is larger than for low frequent items because an often-experienced item is related to several activations of context information.

According to the VT, hits and false alarms are differentially affected by the different classes of words (high and low frequency), which affect the variability of the net input to the nodes in the network. However, on a mathematical level, this effect is independent of what constitutes the item represented by the item features, with associated context information in the context features. It may be an item (a word) or a feature of this word (letters). Never the less, the effect on the variability and thereby on the number of active nodes in the network is the same.

Fame differs from frequency because an encounter with a famous name or face evokes representations specifically related to that person. A certain famous person is associated with specific characteristics, such as a special voice, a face, or a character of a movie – which a non-famous person is not, and these characteristics are reliable. That is, the representation of Anton Chigurh in "No country for old men", or Frank Bullitt in "Bullitt" is relatively unaffected by subsequent encoding of the name Javier Bardem and Steve McQueen, as compared to the analogous effect of repetitive encoding of a non-famous name. Thus, there is a greater match between the pre-experimental context and the context evoked by the study list, because the name is associated with a specific context, for which the subject have a certain pre-experimental knowledge. The pre-experimental context representation of a famous person that is encoded at study is assumed to be similar to the context representation of the same famous name encoded in previous event(s), which we implement by using highly correlated context representations for all contexts associated to a specific famous item. However, the increase in performance in recognition of famous names is not only associated with similar context representations, but also with a high level of distinctiveness. Therefore, repetition of a famous name is beneficial because the context representations are similar, but repetition of a non-famous name is detrimental because the context representations are dissimilar and contextual information associated with one names is easily confused with contextual information for another name. Thus, on an analytical level, the increased recognition performance is modeled by using identical context representations. This was also done because we wanted to avoid adding an additional parameter describing the degree of correlation, and because this would decrease variability.

At each new exposure to the famous name, the retrieved context information overlaps with the context information retrieved at a previous occurrence to a higher degree than for a non-famous name. Thus, repeated exposure to a famous name strengthens the memory representation of that person which leads to a higher net input to the item and context nodes for both new and old items, thereby leading to an increase in performance.

Item, Source and Associative memory

As described above, old items are recognized as old if the number of active nodes exceeds a familiarity threshold. The distribution of the probability for active features depends on the strength and variability of the net input, and the variability and magnitude of the net inputs determine recognition. In the current study, we apply the model on item, source and associative memory.

Source information is based on the context layer. The Hebbian-learning rule used in the model (Sikström, 2001) modifies the weights between the item and context nodes at encoding. At recognition, the encoded pattern for an item and the context is reinstated as cues. Because retrieval of contextual information depends on both the magnitude and variability of the net input to the context layer, the size and variance of the net input determines source memory performance. Thus, the number of active context nodes determines source memory, whereas item memory depends on activity in both the item and context layers. Therefore, the same response criterion applies for item and source memory, which constrains the model with an additional parameter. Thus, for famous names, where the pre-experimental and experimental context representations are highly correlated, the retrieval of relevant contextual information from the study event is stronger than for a non-famous name. Context information related to famous items can more easily be disentangled from irrelevant pre-experimental context information because the retrieved information is highly specific and strongly associated with the famous name. Further, the highly similar pre-experimental and experimental context representation does not decrease the ability to differentiate two syntactically similar (famous) names, because a famous name is highly distinctive and bears individuating features. This is not the case with a non-famous, especially high frequent name, where the retrieved context information is unspecific and may relate to several common names.

For associative recognition, two items are reinstated as cues at the item layer. Given that these cues induce a net input above the activation threshold, the nodes are activated. The reactivation of the context layer can include information previously related to the item, such as spatial information or information about item pairing.

For famous names, with correlated context representations, there is an increase in net input on both the item (pairing of the two items) and the context layers (contextual information about each item). Due to the increase in active nodes in the context layer, famous names will to a larger extent enable access to episodic context information and lead to better source and associative memory as compared to non-famous names. This occurs because the context layer projects to the item layer. Both high and low frequent items should therefore be affected by fame because this manipulation elevates the number of active nodes, and consequently, the net input in both layers affects recognition accuracy for fame and frequency. Because associative memory (e.g., paired associates memory) is modulated by activity in the item layer (the combination of two items) and the context layer (the combination of associative information related to each name), the same response criterion can be used for both associative and item memory. It must be noted that associative memory refers to paired-associates recognition, with the implication that the solution described above not necessarily can be applied on other forms of associative memory. Thus, the account of associative memory is simplified. We implement associative recognition by representing the pairing of the two items in a combined representation in the item layer, whereas the context layer represents contextual information for both items. In the general discussion, we suggest how pairings of two items can be conducted with two distinct representations of the items, which are associated with each other during associative recognition.

Distribution of net inputs and active nodes

Here we describe how the net input and activation of item and context nodes are influenced by fame and frequency, which delineates how and why variability of net input and performance differ for fame and frequency. Figure 2 and illustrates a normal distribution of the probability density of the net input to features and active features for respective variable in each layer. The panels were generated by fitting the models parameters to empirical data in Study 1.

Net input

Normal distributions of the probability density functions of the net input are plotted in Figure 2, which constitutes 4 panels: one for each layer (item and context) for each variable (fame and frequency). Thus, frequency is delineated in panel A and B, and fame in panel C and D.

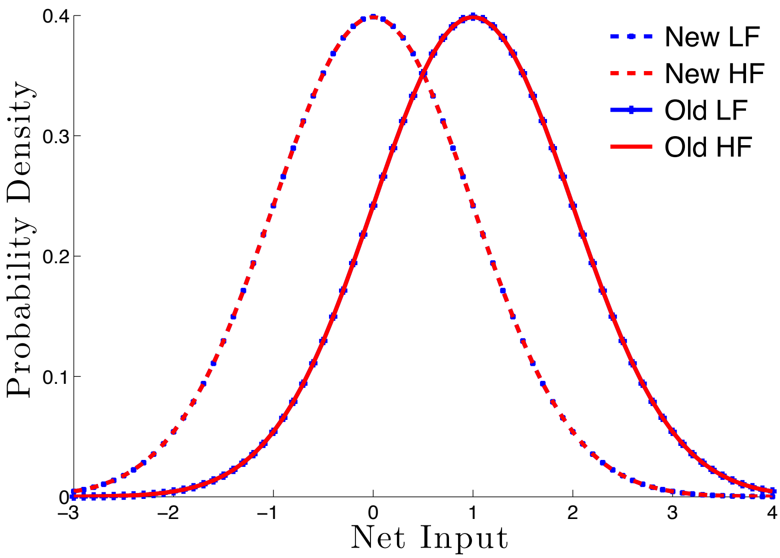
As can be seen in Figure 2, panel A and B, frequency results in an increased variability for high as compared to low frequency items, and a higher net input for old than new items. Theoretically, this would only occur in context layer. For fame (panel C and D), variability is equal for famous and non-famous names, but with a higher net input for famous as compared to non-famous items, similar to the fact that the old items had a higher net input than new items. This effect is found both the item and the context layers.

Active nodes

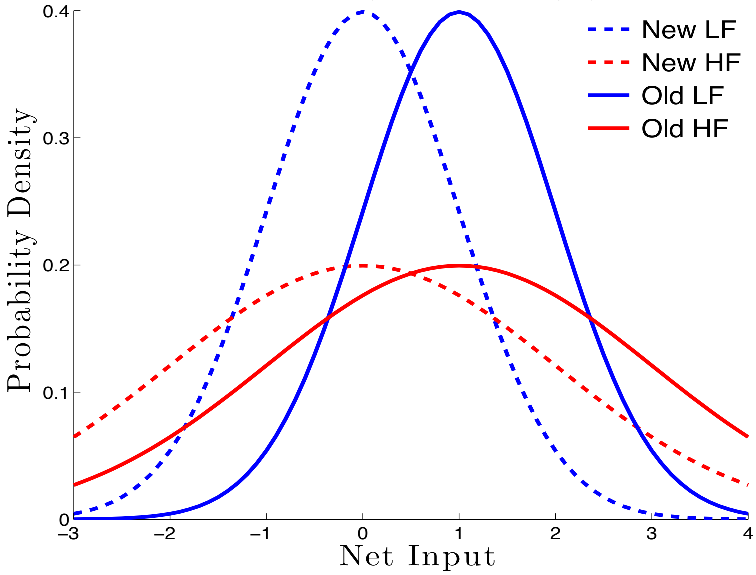
Here we will show that the number of active features is greater for old items than new ones, and for famous as compared to non-famous names. Figure 3, panel A and B, represents the number of active item and context features respectively for the frequency manipulation. As a result of the increased variability for frequency in the net input, and the higher net input for old as compared to new items, a mirror effect occurs in the number of active nodes. These effects are relatively similar in both layers, although with a greater dispersion for the context features. Variability is higher for old than for new items, which is the result of a larger number of active nodes for old items at recognition (but not at encoding).

Figure 2. Net Input to Item and Context layers

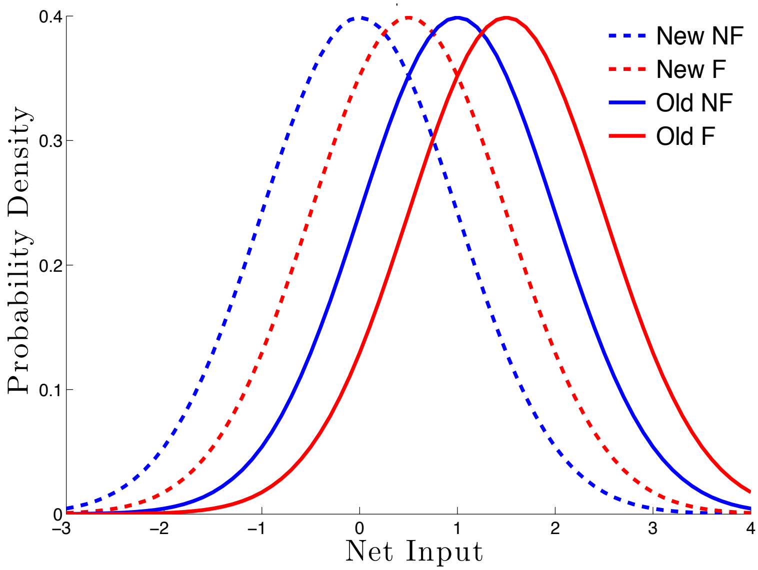
Panel a. Input to Item Features – High and Low Frequency



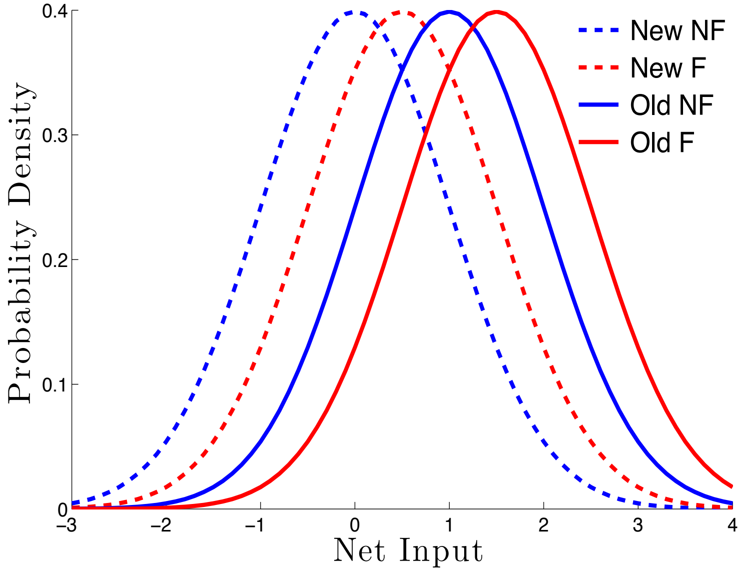
Panel b. Input to Context Features – High and Low Frequency



Panel c. Input to Item Features – Famous and Non-famous



Panel d. Input to Context Features – Famous and Non-famous



Note. Panel a: Input to item features of frequency. Panel b: Input to context features of frequency. Panel c: Input to item features of fame. Panel d: Input to context features of fame. Solid lines represent old items and dotted lines represent new items. HF – High frequency names; LF – Low frequency names; NF – Non-famous names; F – Famous names. The panels were generated by fitting the model to empirical data in Study 1.

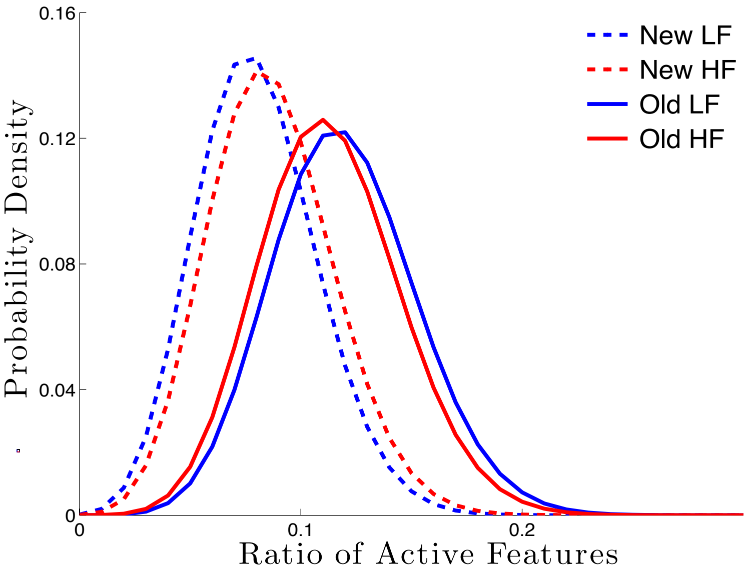
A larger number of active nodes for low compared to high frequency old items accounts for a larger confidence variability (as measured by the z-ROC slope) of low compared to high frequency old items, because an increase in the number of active nodes increases variability in confidence. However, it does not account for a larger variability of new low frequency items compared to new high frequency items, which is a standard finding in the literature (Glanzer & Adams, 1990b; Glanzer, Kisok, & Adams, 1998). This finding is accounted for by introducing variability in the activation threshold, which influences the low frequency net input representation more (due to lower variability) than the high frequency representation (for details, see Sikström, 2004). That is, if there is a higher variability in the activation threshold, the variability in net input increases relatively more for low frequent than high frequent items, because low frequent items has a relatively low variability compared to high frequent items. However, in this paper we do not introduce variability in the activation threshold, simply because this would introduce an additional parameter that would be redundant as we do not present empirical data on the variability in confidence (i.e., ROC data).

For old famous items, the variability in active item features (Figure 3, panel C) is lower compared to new items. Famous, as compared to non-famous names, is associated with a higher variability. Active context features have a similar pattern of distributions, but with a strikingly lower variability and number of active nodes for new, non-famous items, and with completely aligned distributions for old, famous and non-famous, as well as new famous items.

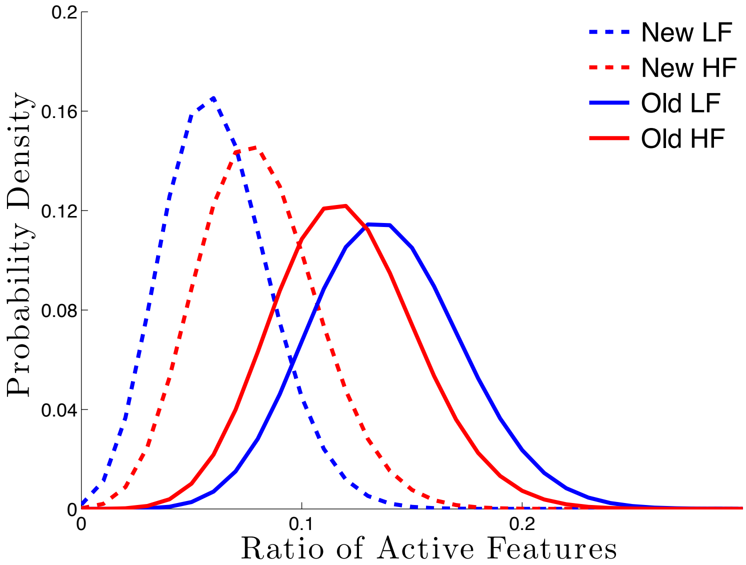
This illustrates that variability in the context layer relate to frequency whereas activity in both the item and context layer relates to fame, and that the stimulus classes influences recognition differently. High frequency leads to a greater variability in the context layer, which decreases the proportion of active features for old items and increases the proportion of active features for new items. Fame, on the other hand, does not significantly influence the variability in the item layer, but increases the proportion of active features in both layers. Consequently, low frequent names, and famous names, are better remembered than high frequent, and non-famous ones.

Figure 3. Active Item and Context Features in Study 1

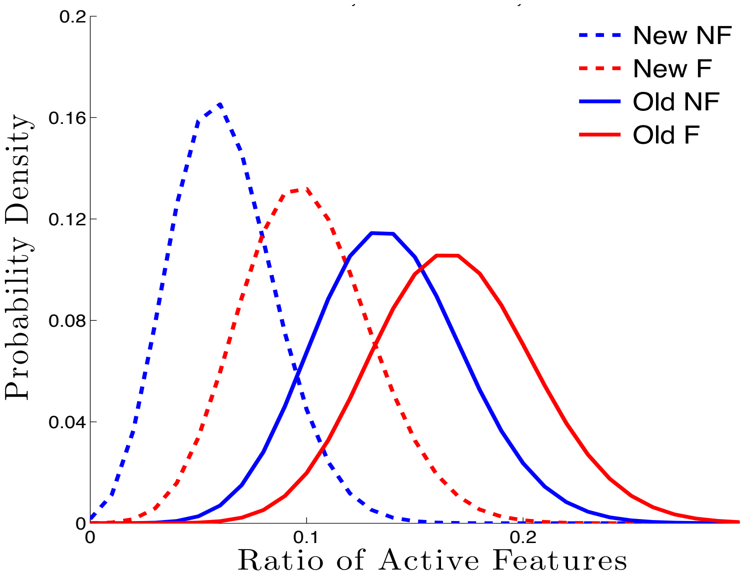
Panel a. Active Item Features – High and Low Frequency



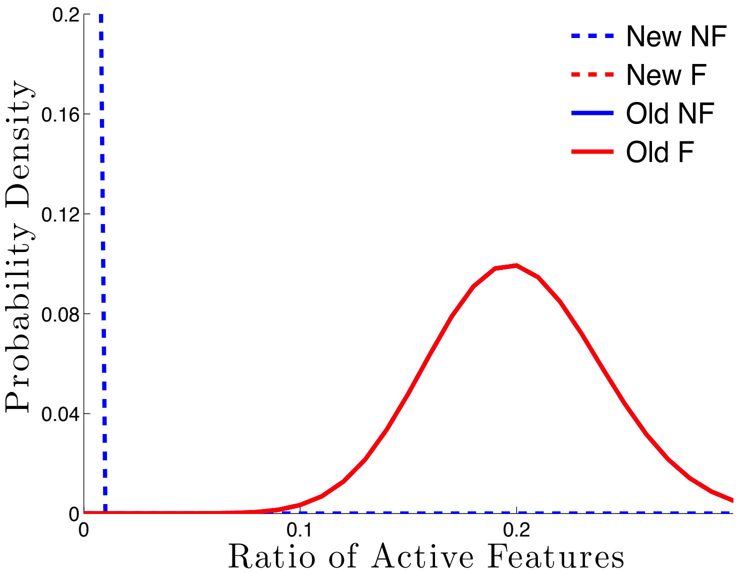
Panel b. Active Context Features – High and Low Frequency



Panel c. Active Item Features – Famous and Non-famous



Panel d. Active Context Features – Famous and Non-famous



Note. Figure 3(a): Active item features for frequency. Figure 3(b): Active context features for frequency. Figure 3(c): Active item features for fame. Figure 3(d): Active context features for fame. Solid lines represent old items and dotted lines represent new items. NF –Non-famous names; F – Famous names; LF – Low frequent names; HF – High frequent names.

Single and Dual Process Models

To assess whether the VT can provide a novel and important model of recognition memory, it will be discussed in the context of the Unequal-Variance Signal Detection model (the UVSD) and the Dual-Process Signal Detection model (the DPSD, Yonelinas, 1994).

To account for variability in memory strength in recognition memory tests, the UVSD and the DPSD is commonly used. In the former, memory is described by a single memory strength variable with different variability in the underlying familiarity distributions that can account for both item memory and confidence data. In the latter (Yonelinas, 1994, 2001), recognition memory is based on two different memory processes: familiarity and recollection. Recognition by recollection, by which contextual information about the study event is retrieved, results in correct high confident hits, which increases the variance of the old item distribution. Thus, a threshold process describes recollection. Recognition by familiarity, on the other hand, results in equal variability for old and new items. Familiarity, then, constitutes a familiarity estimate of the presented item that varies in confidence and accuracy, and is described by a signal detection process.

The distinction between familiarity and recollection have been investigated during several decades, but is still a topic under scrutiny and debate (Dunn, 2004; 2008 Pratte & Rouder, 2012; Rugg & Curran, 2007; Skinner & Fernandes, 2007; Smith & Duncan, 2004; Squire, Wixted, & Clark, 2007; Wixted, 2007; Yonelinas & Parks, 2007) , and in this debate, the UVSD and the DPSD model are commonly compared. Even though both the DPSD and the UVSD has been used frequently, no consensus has been reached as to which one offers the best account of recognition memory (DeCarlo, 2002; Onyper, Zhang, & Howard, 2010).

There are some rather important limitations with both SDT and the DPSD. First, SDT does not provide a detailed account of changes in the variance of the latent distributions, and must rely on the verbally formulated hypothesis of encoding variability (Wixted, 2007) to explain why the old item variance is lower than that of new items. However, the account of EVH is undetailed and debated (Jang, Mickes, & Wixted, 2012; Koen & Yonelinas, 2010, 2013; Starns, Rotello, & Ratcliff, 2012). Further, conventional SDT omits the possibility for dual processes in recognition memory because recognition is described with a single variable of memory strength. Even though the need for an additional process is still debated, there is strong support for this assumption (Diana & Ranganath, 2011; Diana et al., 2006; Rugg & Curran, 2007; Rugg & Yonelinas, 2003; Skinner & Fernandes, 2007; Yonelinas et al., 2010; Yonelinas & Parks, 2007). However, even though the DPSD include a recollection component, the model is dependent on debated assumptions; that recognition memory constitute two qualitatively different retrieval processes, and that recollection is a threshold process (Dunn, 2004; Mickes, Wais, & Wixted, 2009; Ratcliff, Van Zandt, & McKoon, 1995; Slotnick, 2010; Wais, Mickes, & Wixted, 2008). Thus, either one process contributes to recognition, where the variability for old and new items differ because items differ in memory strength (encoding variability), or two processes influences recognition, where the differences in old and new item variability occurs because one process (recollection) elevates performance and the other process (familiarity) leads to variable performance. That the assumptions needed to account for recognition memory with SDT is problematic has been described elsewhere (Koen & Yonelinas, 2010; 2013), as has the assumptions of the DPSD (Glanzer et al., 2004; Qin, Raye, Johnson, & Mitchell, 2001; Slotnick, 2010; Slotnick et al., 2000; Mickes, Wais, & Wixted, 2009). To overcome these limitations, we suggest a parsimonious, formal account of both frequency and fame effects on episodic memory, namely the variance theory.

Aims and Predictions

The main aim of the current study is to provide a more detailed account of the name paradigm with the variance theory. By extending the accountability of the VT to data collected with the name paradigm, a novel account is provided for frequency and fame data, including item, source and associative memory.

Our general predictions are that fame is associated with an increase in both item and source memory (both studies), as compared to non-famous names (controlling for frequency). Also, fame will contribute to successful differentiation of correct and incorrect associative recognition (Study 2). Frequency, on the other hand, is expected to contribute to item memory but not source memory (both studies). Further, higher frequency will give rise to more erroneous associative responses than fame (Study 2).

According to the VT, these predictions can be accounted for by differences in variability and net input to the context and item features mediated by frequency and fame, respectively, as described above. To test this, the model will be fitted to item and source memory in both studies, and be extended to associative memory in Study 2.

Study 1

In the first study, we use the Source Monitoring Framework (Johnson, Hashtroudi, & Lindsay, 1993; Mitchell & Johnson, 2000). When remembering a recent or remotely experienced event, different facets of that experience are bound together during encoding. At test, the participant is instructed to distinguish old from new items (item memory), and when an item is judged old, to determine what context it was presented in during encoding (Johnson, 2006). Thus, remembering a source of an item demands retrieval of contextual information that links the item to its source, whereas item memory does not tax episodic retrieval to the same degree. Thus, the source memory framework can be used to investigate whether the variables frequency and fame have differential effects on these two memory performances (item and source memory).

The VT predicts that fame and frequency affects recognition performance differently, because fame increases the number of active features for the retrieved name with no effect on the variability of the net input. Frequency, and foremost common names, introduces increased variability in the net input and lowers performance. Consequently, source memory will be affected by fame and not frequency, because retrieval of contextual information necessary for accurate source memory judgments demands a higher number of active nodes, rendered by the highly correlated context representations. We also want to investigate whether the two variables frequency and fame have different effects on hits and false alarms. We will fit the variance theory to correct and incorrect responses for item and source memory.

Participants

The experiment took place on two occasions. Thirty-two undergraduate students (23 women) with a mean age ± SD of 23.59 ± 3.96 years (range 20-36 years) were recruited from Kristianstad University (Experiment 1). Subjects received a lunch voucher issued by the campus restaurant (worth approx. 10 $) for participation. Another twenty-eight undergraduate students (18 women) with a mean age of ± SD of 25.21 ± 4.36 years (range 20-34 years) were recruited from Lund University (Experiment 2). The two data sets were analyzed together because the only difference between them was elapsed time between collections. Lab was therefore included as a between-subject variable in the analysis. There was no interaction with this factor.

Materials

The set of names used in this experiments was a revised version of the pool used in previous experiments (Stenberg, Hellman, & Johansson, 2008; Stenberg et al., 2009), updated along both the fame and the frequency dimensions prior to data collection. Frequency was operationalized as the number of hits in the Swedish, nation-wide telephone directory (www.eniro.se). It was log transformed and dichotomized into frequent and infrequent, used as the variable frequency. Fame was operationalized as the number of Google hits in the Web editions of four national newspapers (www.dn.se, www.svd.se, www.expressen.se, www.aftonbladet.se) and two television networks (www.svt.se, www.tv4.se). Summed hits were log transformed and dichotomized into famous and non-famous, used as the variable fame. Searches for famous names were made by a Visual Basic program, using a program interface published by Google (www.google.com). To the current experiments, 96 names were randomly selected from the pool. That is, 24 of each name class: famous, common ones (e.g. Tom Jones), famous, uncommon ones (e.g. Javier Bardem), non-famous, common ones (e.g. John Smith) and non-famous, uncommon ones (e.g. Sebastian Weisdorf). The updated material was validated by participant ratings on the level of fame and frequency for each name at the end of Experiment. This showed a correlation between our a priori ratings (internet searches) and the participant ratings for fame (0.91) and frequency (0.95). Further, the two dimensions were unrelated. Internet ratings of frequency were uncorrelated with participant ratings of fame (0.04) and Internet ratings of fame were uncorrelated with participant ratings of frequency (0.17).

Procedure

Participants were tested in groups of 6 – 14 (Experiment 1), and 3 – 6 (Experiment 2) at a time in a laboratory, seated in separate booths, each with a computer on which the experiment was run, using e-prime. They were given verbal instructions at the outset, and written instructions during the experiment. In two counterbalanced study-test blocks, each participant was presented with 24 names per block (6 per name class) during study and 48 names at test (12 per name class), half of which were new.

At study, each name was presented during 2 sec., in green or red color, and the task was to remember both the name and the color it was presented in for a subsequent memory test. For all participants, names were randomly assigned regarding choice of names, presentation order and stimulus color. At test, participants were presented with previously studied names intermixed with an equal number of new ones, in black color. The task was to differentiate old from new names, and, if accepted as old, to determine whether the name was presented in green or red color during study.

Results and Discussion

We calculated d` (recognition memory performance) and C (the criterion) in accordance with Snodgrass and Corwin (Snodgrass & Corwin, 1988), as well as hits, false alarms and source memory performance. For item memory, correct old judgments were collapsed across source memory. These five measures were included in repeated measures ANOVAs.

As can be seen in Table 1, where averages are reported collapsed across experiments, frequency and fame had different effects on recognition accuracy, as indexed by d`, hits and false alarms. High, compared to low frequency, decreases hit rates but increases false alarms. Thus, our results are in accordance with the mirror effect of recognition memory.

Item Memory

Turning to the statistical analysis, we tested the effects of frequency and fame in a 2 (frequency) * 2 (fame) repeated measures ANOVA with lab (Experiment) as between subject variable, on d´, the criterion (C), hit rates, false alarms and source memory performance. Starting with d`, the effect of frequency was large [$F(1,58)=43.02$, $p<0.001$, $eta^2=0.46$], however weaker than the effect of fame [$F(1,58)=78.51$,

$p<0.001$, eta^2=0.57], and with no interaction. Thus, infrequent names were better remembered than frequent ones, and famous names exceeded non-famous ones. Fame had a large impact on hit rates [$F(1,58)=122,67$, $p<0,001$, eta^2=0.68] as did frequency [$F(1,58)=51.70$, $p<0.001$, eta^2=0.47]. There was an interaction [$F(1,58)=4.19$, $p<0.05$, eta^2=0.07], because of lower performance for the non famous, high frequent names. Frequency contributed to false alarms [$F(1,58)=27.42$, $p<0.001$, eta^2=0.32], as did fame, but with a smaller effect [$F(1,58)=4.18$, $p<0.05$, eta^2=0.06], and with no interaction. There was a large effect of fame, and fame only, on the criterion [$F(1,58)=43.72$, $p<0.001$, eta^2=0.43)[1].

Table 1. Mean performance (St. D) for item and source memory in Study 1

Name class		d`	C	H	FA	SMP
Famous	*Frequent*	2.02 (0.79)	-0.02 (0.44)	0.86 (0.16)	0.14 (0.18)	0.73 (0.23)
	Infrequent	2.58 (0.57)	-0.04 (0.26)	0.95 (0.07)	0.08 (0.15)	0.73 (0.22)
Non-famous	*Frequent*	1.39 (0.91)	0.42 (0.49)	0.55 (0.25)	0.12 (0.13)	0.58 (0.28)
	Infrequent	1.87 (0.77)	0.39 (0.44)	0.71 (0.22)	0.03 (0.06)	0.56 (0.24)

Note. d`, C (the criterion), Hits and False Alarms (denoted FA) are calculated for item memory (correct old/new judgment irrespective of source accuracy). d` and C was calculated in accordance with Snodgrass and Corwin (1988). SMP represent source memory performance, calculated as (correct source response)/(incorrect source response + correct source response). Reported values are collapsed across Experiment 1 and 2.

Source Memory

Source memory performance (SMP) is presented in Table 1, which shows that fame leads to better source memory than frequency. In a 2 (frequency) * 2 (fame) repeated measures ANOVA, with experiment as between subject variable, fame had a reliable effect [$F(1,58)=26.40$, $p<0.001$, eta^2=0.31], whereas frequency did not [$F(1,58)=0.07$, *ns*.].

Model Fitting

The model was fitted to hits, false alarms and source memory responses for the four different name classes (famous, high frequent/ famous, low frequent; non-famous, high frequent and non-famous, low frequent) with 12 data points (6 trials/name class over 2 study/test blocks), using maximum likelihood estimation (MLE). The model was fitted with five parameters: the standard deviation of the net input $\sigma_h(f)$ for high and low frequency, the expected net input to famous items $\mu_h(F)$, the standard deviation of input to the item layer $\sigma_h(I)$ and the activation threshold (T). The model was fitted to both group and subject data.

The residuals for predicted and observed values for hits, false alarms and SMP for the group fit are reported in Table 2. For item and source memory, the fit reached an MLE of 341.66. Values for the predicted item and context layer activity can be seen in Table 2. As can be seen, activity in the context layer for old names is inversely related to frequency, whereas activity increases for new high frequent as compared to new low frequent names. Also, activity in the context layer is higher for famous than non-famous names. Item layer activity increases with fame, but is unaffected by frequency.

The individual fit resulted in an average MLE of 453.45, with a standard deviation of 196.49. The fitted parameters had an average of 0.13 and 0.65 for the standard deviation in net input for low and high frequency $\sigma_h(f)$, and 0.038 for net input for famous items $\mu_h(F)$. The standard deviation of the net input to the item layer $\sigma_h(I)$ equaled 0.15 and (T) reached 0.74.

In the first study we confirmed our hypothesis regarding the effects of the variables frequency and fame on episodic recognition memory, with an effect of both variables on item memory, but only of fame on source memory. Fame had a greater effect than frequency on hits, and frequency had a greater effect on false alarms as compared to fame. This pattern of effects is in accordance with those in previous experiments with the name-paradigm, where the effect of frequency and fame on false alarms and hits were dissociated (Stenberg, Hellman & Johansson, 2008; Stenberg et al., 2009). Famous names were retained better than non-famous ones, thus, the prior experience of these names strengthens memory. However, frequency, reflecting prior experience as well, has the opposite effect.

To understand this difference, that is, different performance in source memory, hits and false alarms, we suggest the VT. Accordingly, the two types of pre-experimental familiarity have differential effects on recognition memory because high frequency items introduce variability in the network whereas fame does not. Common names are related to a high number of contexts. Therefore, at test, the reactivation induces a high variability in the context layer when the activation pattern from the pre-experimental study phase is reinstated.

Table 2. Residuals for predicted and observed responses and predicted item and context layer activity in Study 1

Name class		Residuals			Item layer		Context layer	
		Hits	FA	SMP	New	Old	New	Old
Famous	*Frequent*	0.071	-0.036	-0.04	0.196	0.769	0.156	0.808
	Infrequent	0.065	-0.018	-0.04	0.196	0.769	0.000	1.000
Non-famous	*Frequent*	-0.037	0.054	0.001	0.082	0.579	0.050	0.594
	Infrequent	-0.001	-0.011	-0.02	0.082	0.579	0.000	0.842

Note. Residuals reflect difference between observed and predicted data, where a negative value reflects a lower predicted value than the observed. SMP represent source memory performance, calculated as (correct source response)/(incorrect source response + correct source response).

Due to the high variability in the underlying distribution, the participant has difficulties distinguishing old from new high frequent names, with more erroneously responses as a result. For uncommon names, there are fewer contexts pre-experimentally associated with the name, and therefore, the variability in net input is considerably lower as compared to high frequent names. Famous names are associated with similar pre-experimental contexts and therefore induce a higher net input to the item and context features with unaffected variability (see Figures 2 and 3). The low variability in input to the context and item layers increases the proportion of active features for old items. Consequently, low frequent names, and famous names, are better remembered than high frequent, and non-famous ones. Frequency, which affected activity in the context layer, was higher for rare than common names. Fame affected the item layer to a higher degree than the context layer. Here, activity was higher for famous than non-famous names. Fame increased performance for old items as compared to non-famous names, and also increased the proportion of incorrectly endorsed new items for famous, high frequent, as compared to famous low frequent names.

Study 2

We conducted a second experiment to replicate Study 1, and to extend the accountability of VT. We constructed an associative recognition test with the name-paradigm using a paired associate test. Here, items are presented as pairs at study, and at test, pairs are presented as intact combinations, as completely new pairs or as recombined pairs (a combination of a new and an old item). Further, the name pairs were presented in 4 different spatial positions at study, which functioned as contextual information in a source memory judgment at test. Thus, participants made both judgments about item-item associations and item-context associations. To compare associative memory for famous and non-famous names with varying frequency, we will test differences in both endorsement and rejection of old items and recombined items, because a correct rejection of a recombined item necessitates episodic retrieval of context information related to the specific item, which is necessary for associative recognition. Data is collected using the paired-associate paradigm, meaning that associative memory is defined as retrieval of contextual information about the encoded item, not as retrieval of additional item information (i.e., pairing of two qualitatively different items such as a name and a face).

We predict that both fame and frequency will increase correct item-item judgments, and that only fame will contribute to successful item-context judgments (i.e., processing of retrieved context information). Frequency will contribute more to false alarms than fame, and more to incorrect item-context judgments as compared to fame (because frequency is associated with less retrieval of context information). Also, fame will increase correct source memory judgments, whereas frequency will not. As in Study 1, item and source memory will be fitted to the VT, and in the current experiment, this will be extended to associative memory, by fitting correct and incorrect old and new responses to recombined items.

Participants and Material

Thirty students (22 women) with a mean age ± SD of 25.23.x ± 4.26x years (range 19-36 years) were recruited from psychology courses at the university of Lund. To the current experiment, 192 names were collected from the pool used in Study 1, half of which were female.

Procedure

Participants were tested in groups of 2 – 6 at a time in a laboratory, seated in separate booths, each with a computer on which the experiment was run, using e-prime. They were given verbal instructions at the outset, and written instructions during the experiment. At study, each participant was presented with 12 name combinations, each combination consisting of 2 names from the same name class. At test, the participant was presented with 24 name combinations (6 combinations per name class). Of these, 8 were intact old combinations (2 per name class), 8 were recombined combinations (one studied combination generated 2 recombined ones, 2 per name class) and 8 were completely new combinations (also, 2 per name class). In total, 192 names were used in 4 counterbalanced study-test list cycles.

At study, each name combination was presented at one of four positions (up to the left, up to the right, down to the left or down to the right), each location separated with a black frame. Presentation of the name combination was preceded with a fixation cross for 500 ms., located at the same position as the forthcoming name combination (spatial location was equally assigned to every test phase). The task was to remember both the name combination and the spatial location it was presented in for a subsequent memory test. To enhance encoding, participants accomplished an orienting task. For each presented name combination during study, they stated the gender or the combination (both male, both female or different sex). The material was therefore divided into equal groups of male and female numbers. Each item was presented for 2,5 sec., and the orienting task was administered in a new window and was self-paced. For all participants, names were randomly assigned regarding choice of names and combination, providing they matched the gender criteria. Further, each combination was randomly assigned regarding study list, presentation order and spatial location during study.

At test, each name combination was presented at the center of the screen. Participants were instructed to judge each combination as old, new or recombined, at a self-paced rate. If they responded old, the current name combination was presented centered on the screen with the 4 spatial locations boxes in respective corner, also at a self-paced rate. Each box contained a number, which corresponded to a numerical key on the keyboard, used as source memory response. To maintain maximal study-test overlap, the old names of the recombination's as well as the intact combination were presented in the same order, and sex –assignment was held constant (i.e., using English equivalents, if combo *John Smith – Jane Cooper* was studied, the recombination constituted *John Smith – Linda Johnson*, and *Robert Brown – Jane Cooper*).

Results and Discussion

We calculated d` and the related bias measure C, as well as hits and false alarms. For item memory, correct old judgments were collapsed across source memory. We also calculated source memory performance, correct rejections and false alarms for recombined items (i.e., recombined response to recombined items and old response to recombined items). These seven measures were included in repeated measures ANOVAs.

As can be seen in Table 3, there are differential effects of, foremost, frequency, on recognition. Low frequency names are associated with higher performance (d′) than high frequency names. Famous names have a higher performance than non-famous ones.

Table 3. Mean performance (St. D) for item, associative and source memory for Study 2

Name class		d'	C	Hit item	FA item	CR rec.	FA rec.	SMP
Famous	*Frequent*	1.63 (0.62)	-0.38 (0.32)	0.67 (0.18)	0.09 (0.10)	0.62 (0.28)	0.19 (0.21)	0.62 (0.25)
	Infrequent	2.64 (0.61)	-0.21 (0.33)	0.88 (0.21)	0.01 (0.04)	0.58 (0.26)	0.12 (0.13)	0.76 (0.26)
Non-famous	*Frequent*	1.56 (0.77)	-0.67 (0.42)	0.53 (0.27)	0.03 (0.07)	0.43 (0.29)	0.12 (0.10)	0.56 (0.35)
	Infrequent	2.10 (0.68)	0.48 (0.31)	0.71 (0.21)	0.01 (0.04)	0.49 (0.26)	0.12 (0.11)	0.55 (0.23)

Note. FA Rec. represents False Alarms for recombined items (incorrectly accepting a recombined item as old) and CR Rec. represents correct classification of recombined items (i.e., FA rec. and CR rec. reflect associative memory).

Item Memory

Turning to the statistical analysis, we tested the effects of frequency and fame in a 2 (frequency) * 2 (fame) repeated measures ANOVA, on the eight measures. Both frequency [$F(1,30)=55.25$, $p<0.001$, eta^2 =0.65], and fame [$F(1,30)=6.60$, $p<0.05$, eta^2 =0.18] had an effect on d`. Thus, infrequent names were better remembered than frequent ones, and famous names were better recognized than non-famous ones. There was an interaction [$F(1,30)=6.76$, $p<0.05$, eta^2 =0.18], because high frequency decreased performance. Further, frequency [$F(1,30)=35.9$, $p<0.001$, eta^2 =0.54] and fame [$F(1,30)=20.0$, $p<0.001$, eta^2 =0.40] contributed to hits. False alarms were affected more by frequency [$F(1,30)=13.46$, $p<0.01$, eta^2 =0.31] than by fame [$F(1,30)=6.4$, $p<0.05$, eta^2 =0.17]. There was a fame*frequency interaction for false alarms [$F(1,30)=10.7$, $p<0.005$, eta^2 =0.26], reflecting more erroneously responses at higher levels of frequency. Participants applied a rather liberal response criterion, with an effect of both frequency [$F(1,30)=9.37$), $p<0.01$, $eta^2=0.24$], and fame [$F(1,30)=24.51$, $p<0.001$, $eta^2=0.45$][2]. Thus, participants were more conservative in endorsement of non-famous items, and more liberal for low frequent items.

Associative Memory

Only frequency contributed to false alarms for recombined items [$F(1,30)=4.66$, $p<0.05$, $eta^2=0.13$], whereas only fame contributed to correct rejections for recombined items [$F(30)=21.55$, $p<0.001$, $eta^2=0.42$]. There was no reliable interaction.

Source Memory

For source memory performance (SMP), fame had an effect [$F(1,30)=14.06$, $p<0.005$, $eta^2=0.32$], whereas frequency did not [$F(1,30)=2.72$, *ns.*]. There was no sign of an interaction.

Model Fitting

As in Study 1, the model was fitted to hits, false alarms and source memory responses for the four different name classes (famous, high frequent/ famous, low frequent; non-famous, high frequent and non-famous, low frequent) with 12 data points, using MLE for 30 participants. Further, the model was fitted to correct and incorrect responses to recombined items (fitted with the same parameters as in Study 1, for N=30 and 8 data points). We choose to fit item and source memory (12 data points), and associative memory (8 data points) in two separate fits because it unlikely that associative recognition is determined by the same criterion setting as item and source memory. The model was fitted on group and individual level.

In the first fit (item and source responses), on group level, the standard deviation of the net input for high and low frequency $\sigma_h(f)$ was 0.66 and 0.05, respectively. The expected net input to famous items $\mu_h(F)$ equaled 0.17 and $\sigma_h(I)$ was 0.44. The activation threshold (*T*) reached 0.94. Residuals for predicted and observed values for item, source and associative memory for group data are reported in Table 4. The fit reached an MLE of 27.33.

In the individual fit for item and source memory responses (12 data points), the fitted parameters averaged 0.27 and 0.16 for the standard deviation of the net input for high and low frequency $\sigma_h(f)$, 0.18 for input for famous items $\mu_h(F)$ and 0.63 for

$\sigma_h(I)$. The activation threshold (T) reached 0.77. The average MLE was 102.61 with a standard deviation of 24.35.

Table 4. Residuals for Predicted and Observed responses for Study 2

Name class		Hits	FA	SMP	Hits Rec.	FA Rec.
Famous	*Frequent*	-0.001	0.016	-0.08	0.017	0.033
	Infrequent	0.028	-0.07	0.05	-0.027	-0.006
Non-famous	*Frequent*	-0.011	-0.011	0.01	-0.083	-0.030
	Infrequent	0.000	0.004	0.00	-0.024	0.003

Note. Residuals reflect the difference between observed and predicted data, where a negative value reflects a lower predicted value than observed. SMP represent source memory performance.

Frequency and fame was related to activity in the context and item layers respectively. Values are reported in Table 5, for item, source and associative memory. As can be seen, for item and source memory, context layer activity was higher for low than high frequent names and for famous as compared to non-famous names. Activity was greater for new frequent than for new low frequent names. Activity in the item layer was affected by fame, but not frequency, and activity in the context layer was more affected by frequency than fame.

In the second fit, associative memory, the model was fitted to endorsement of intact items and recombined items. Values for $\sigma_h(f)$ was lower for low (1.35) than high (1.87) frequent items, and $\sigma_h(I)$ reached 0.182. The value for $\mu_h(F)$ was 0.078, and the activation threshold reached 0.99. The fit generated an MLE of 34.74. Predicted values for the item and the context layers revealed that the item layer was affected by fame, but not frequency. Incorrect endorsement of new items was higher based on context layer activity than by activity in the item layer, and also, higher for common than rare names. Thus, context layer activity was more affected by frequency than fame, as compared to activity in the item layer.

Table 5. Context and Item Layer Values for Study 2

Name class		Item and Source Memory				Associative Memory			
		Item		Context		Item		Context	
		New	Old	New	Old	New	Old	New	Old
Famous	*Frequent*	0.03	0.70	0.11	0.64	0.00	0.67	0.31	0.52
	Infrequent	0.03	0.70	0.00	1.00	0.00	0.67	0.25	0.53
Non-famous	*Frequent*	0.13	0.55	0.07	0.53	0.00	0.52	0.30	0.50
	Infrequent	0.01	0.55	0.00	0.87	0.00	0.52	0.23	0.50

In the individual fit, the following parameter values were reached. The standard deviation of net input $\sigma_h(f)$ was 0.31 and 1.01 for low and high frequency, respectively. The net input for famous items $\mu_h(F)$ averaged 0.46 and $\sigma_h(I)$ was 1.39. The activation threshold was 0.48. This generated an average MLE of 166.47.

In the current experiment, both low frequency and fame contributed to recognition memory performance and hits, whereas frequency increased both false alarms for item memory and associative memory. Fame had no effect on recombined false alarms, and, alone, contributed to source memory performance and correct rejections for recombined items, with no effect of frequency. As described in Study 1, the VT suggests how frequency influences the variability in net input to the context layer and how fame affects the net input to the item layer in the model, which consequently affects the number of active features. For item and source memory, the VT provided a reasonable good fit with a reasonably small error between predicted and empirical values. Predicted values for the item and context layers showed similar parameter values and similar fits to empirical data as in Study 1. Fame increased the net input to the item layer whereas frequency contributed to the context layer. For associative memory the effects of fame and frequency on the item and context features are differentiated.

Results from both Study 1 and 2 support the prediction that fame and frequency are selectively related to activity in the item and context layers, respectively. Performance for source and associative memory was considerably higher for famous as compared to non-famous names, whereas both fame and frequency affected item memory. When making recognition judgments about a recombination of previously studied and unstudied names (i.e., associative memory) correct responses necessitate retrieval of episodic contextual information about the status of the item pairings during study (Yonelinas, 2002). But, when this recombination encompasses non-famous names, for which retrieval of contextual information is lower, performance suffers. Instead, the name combination is accepted as old due to an elevated familiarity signal, based on an incremental process (Mandler, 1980).

There was a difference in the response criterion (C) adapted by the participant over the two experiments. In the first experiment, fame alone had an effect on the criterion, reflecting more liberal recognition decisions for famous items as compared to non-famous items. In Study 2, both variables contributed to this measure, although with a larger effect of fame. Participants were more conservative in their memory decisions for the non-famous names and more liberal for low frequent names.

General Discussion

In two experiments, we have applied the variance theory on recognition memory data, encompassing item memory, source memory and associative memory. Using the name paradigm, we have recorded data for Swedish personal names varying in fame and frequency. Thus, we have compared memory for high and low frequency for famous and non-famous names. In both Study 1 and 2, frequency affected the variability of the net input to the context layer of the model. This occurs because high frequent names are associated with several pre-experimental contexts, making the item less specific as compared to a low frequent name, which is associated with a low number of contexts. Therefore, a rare name increases the number of active features, leading to higher memory performance, as compared to a common name. Variability is lower for new than for old items, and decreasing for low frequent, new names. Thus, in the VT, variations in performance for high and low frequency are accounted for by differences in variability for the two levels of frequency (which is related to different magnitude of pre-experimental familiarity) in the input to the context layer.

The difference in performance for famous and non-famous names (when frequency is controlled) is explained differently. Here, the variability for famous and non-famous names is equal, but a higher net input to both the item and the context layers occurs for famous names. This occurs because famous names, as compared to non-famous names with the same frequency, are associated with similar contexts and higher specificity, resulting in greater match between the pre-experimental contexts and the contexts induced in the study phase of the experiment. Because there is equal variability in input to both layers the proportion of active features increases for famous as compared to non-famous names in the item layer. In the context layer, old famous and non-famous as well as new famous names has equal variability and leads to the same number of active features. New, non-famous names result in lower net input and a lower number of active features.

We argue that frequency and fame reflect two different types of semantic knowledge, and that this difference is important to understand the differential effect on recognition memory. Frequency, on the one hand, is a generic knowledge about statistical regularities of the world, such as how common or uncommon a name is. Within the framework of the VT, this is understood as that there is a greater variability in net input for high frequent names, which makes the connection between the item and context information less specific. Fame, on the other hand, evokes a multitude of associations about the presented name, and reinforces the encoding by using previous knowledge about the name-bearer. Thus, at test, the set of associations evoked at study are likely to be evoked once again. Famous names have a strong and distinctive association between context and item, where the pre-experimental memory representation is similar to that evoked in the experimental situation. Therefore, the familiarity induced by the study phase increases the connection between the presented name and the associated context, but does not introduce an increment in variability in net input to the context layer.

The beneficial effect of fame on episodic recognition may also be understood by relating to a common finding in the memory literature, namely the encoding specificity principle. Accordingly, memory performance is highly influenced by consistency of contextual information of the item at encoding and retrieval (Tulving & Thompson, 1973). Because the context information evoked by presentation of a famous name seems to provide stable and autobiographically significant associations (Westmacott & Moscovitch, 2003), the contextual information at encoding is highly related to that at a previous occurrence, which strengthens these specific representations for subsequent retrieval. The encoding specificity principle also indicates why frequency introduces more recognition errors than fame, as the associate contextual information for a high frequent item is difficult to specifically relate to the presented item.

The VT was fitted to empirical data in Study 1 and 2, revealing that the model accounted for a majority of the data given the overall fit and predicted response probabilities. However, to confirm that the result of the fit match the empirical data, we compared predicted responses with confidence intervals for the sample for hits and false alarms, which are reported in Table 6. As can be seen, the majority of the predicted values are encompassed by the 95% confidence intervals, with a few exceptions. In Study 1, for famous, infrequent names, there was a slight deviation for false alarms. Further, both frequent and infrequent non-famous names deviated from the confidence intervals for false alarms. The former predicted response probability was a bit below the confidence interval, whereas the latter just exceeded it. In Study 2, all predicted values were encompassed by the confidence intervals for the sample means.

Table 6. Predicted H, FA and confidence intervals for Study 1 and 2

Name class		Hits				False Alarms			
		Study 1		Study 2		Study 1		Study 2	
		Pr.	CI	Pr.	CI	Pr.	CI	Pr.	CI
Famous	*Frequent*	0.84	0.82-0.90	0.67	0.60-0.73	0.14	0.09-0.19	0.08	0.05-0.12
	Infrequent	0.88	0.93-0.97	0.85	0.81-0.96	0.11	0.04-0.12	0.02	-0.00-0.03
Non-famous	*Frequent*	0.58	0.48-0.61	0.54	0.43-0.63	0.05	0.08-0.15	0.04	0.00-0.06
	Infrequent	0.71	0.66-0.77	0.71	0.63-0.78	0.04	0.01-0.04	0.01	-0.00-0.03

Note. Pred denotes the predicted value of the VT, and CI denotes confidence interval for the sample mean. Predicted values that are not captured by the confidence interval are shaded grey.

In the current implementation of the variance theory, associative memory, as reflected by performance for the paired associates test, is simplified. Item, source and associative memory scores were fitted with the same number and value of parameters and a common response criterion. Thus, the model was not elaborated specifically to account for associative memory, instead, these scores were handled with the assumption that the item pairings were accomplished in the item layer. When presented with a pair of names from the same stimulus class (for example, famous, high frequent names), respective name is represented in the item layer as a pair, and the associated context information (e.g., the voice of a singer and a movie character of an actor) is represented in the context layer. The model does not describe how the context information of two different items interacts in the context layer, and neither, how input to respective layer is affected by the item pairings. Therefore, the model could be extended with three connections within the item layer, where each item is subtracted with the context layer value, and the first item is subtracted with the value of the second item. Then, the criterion used for item memory is adequate to use for paired associates performance.

One or two retrieval processes?

In the recognition memory literature, the debate on whether a recognition response is governed by one or two processes is ubiquitous (Dunn, 2004; Pratte & Rouder, 2012; Rugg & Curran, 2007; Skinner & Fernandes, 2007; Smith & Duncan, 2004; Squire, Wixted, & Clark, 2007; Wixted, 2007; Yonelinas & Parks, 2007; Yonelinas, 2002). The VT has made no suggestion in this matter, because the empirical data has been accounted for independently of the assumption of two processes. However, several review articles provide support for the assumption that recognition memory constitute two processes; one assessing general memory strength and one that enables the re-experience of the study episode (Diana & Ranganath, 2011; Diana et al., 2006; Rugg & Curran, 2007; Rugg & Yonelinas, 2003; Skinner & Fernandes, 2007; Yonelinas et al., 2010; Yonelinas & Parks, 2007). It can also be argued that a limitation with signal detection theory is the lack of a recall process as described by dual-process models, given the support for dual-process theory according to the reviews above.

The VT can encompass a conceptualization of recognition memory decisions where a recall process is distinguished from item retrieval in the absence of contextual information. Access to both the proportion and specificity of contextual episodic information is related to the number of active context and item nodes, which can relate to different levels of episodic retrieval. As high frequency evoke more pre-experimental contexts, variability in the context layer increases and recognition performance decreases. For low frequent names, variability in net input is lower in the context layer, as compared to high frequent names (See Figure 2). This was also suggested by the model fit, where activity in the context layer (i.e., presumably reflecting familiarity) was higher than the activity in the item layer. Familiarity increased for low as compared to high frequency and was higher for famous than non-famous names. This prediction differs from other dual process models (see Yonelinas, 2002 for a review), where high, rather than low frequency is assumed to elevate familiarity. Support for this notion can also be found in previous work with the name-paradigm (Stenberg et al., 2009), where the electrophysiological correlate of familiarity (the FN400 old/new effect) was more pronounced for low than high frequent names. As described in the introduction, high frequency introduces high variability in net input to the context layer, which results in a lower number of active nodes. Therefore, low frequency items are better retained, and seem more familiar than high frequency items at retrieval. This can be related to George Mandler's conceptualization of recognition memory. When the testee is presented with an item that matches the template of the underlying representation for that item in memory, a feeling of familiarity is produced. A common object is frequently and recently exposed to the testee, thus it is better integrated to memory, and therefore is better matched to the item in memory. However, the integration of an item in memory, that is, a strengthening of the memory trace, is always relative the original integration. Expressed differently, an item seldom experienced has a lower integration from the beginning, as compared to a common item. Therefore, at retrieval, it is the relative increase in integration as a result of the recent presentation (at study) that results in higher familiarity for low frequent as compared to high frequent item (Mandler, 1980; 2008).

The context layer provides representations of contextual information related to the test item, and when such information can be retrieved, recollection leads to access to contextual information projected to the item layer. Due to the lowered variability in net input for famous names, the number of active features increases, resulting in higher overall memory performance, and accurate retrieval of contextual information. As demonstrated by the fit in both experiments, activity in the context layer varied with frequency to a higher degree than activity in the item layer, and activity in the item layer varied with fame but not frequency. In Study 2, for associative memory data, there was no activity for new items in the item layer, in accordance with the notion that paired-associate memory is mediated by recollection, inducing a low number of false alarms. In addition, both source memory and paired-associate memory was affected by fame, not frequency.

The relation between the two variables fame/frequency and recollection/familiarity, and the item and context layers suggests a mapping of the assumed retrieval processes and pre-experimental familiarity. Frequency can be related to familiarity because it induces a heightened variability in the context layer (i.e., inputs from the item layer), which has a detrimental effect on recognition performance, including source memory and associative memory. The variability of the context layer increase with frequency because there are a larger number of contexts associated with the item, resulting in heightened variability in the context representation. Fame may be related to recollection because it leads to activation of similar contexts, with unaffected variability in both the item and the context layers. This results in greater performance on item- and source memory and paired associates tests. Here, the variability of the input to respective layer is unaffected, but with an increase in the number of active nodes. Because source- and associative memory performance was positively related to fame, in contrast to frequency, and because the two variables can be selectively related to the item and context layer of the VT, it seems plausible that two different processes contributes to recognition.

However, the possible link between the item and context layers and the assumed retrieval processes recollection and familiarity does not suggest that two independent processes are necessary to account for the empirical data. Squire and colleagues (Squire, Wixted, & Clark, 2007) recently proposed that differences observed in recognition memory performance that often are interpreted as consistent with dual-process theory, rather reflects a difference in weak and strong memory. Nevertheless, further research is needed to establish whether the eventual relation between the item layer and recollection on the one hand, and the context layer and familiarity on the other hand is valid, or even necessary to account for recognition memory data.

Another limitation with signal detection theory is the dependency on the auxiliary hypothesis of encoding variability (Wixted, 2007) to explain why the old item distribution has a higher variability than that of new items. Even though this discussion lies outside the scope of the current article, because we have not recorded confidence responses, it is worth noticing that the VT can inherently account for the empirical difference in variability in the item distribution by relating confidence and response bias to the number of active features (Sikström, 2001, 2004).

Neural network are known to be sensitive to the correlated inputs, where similar inputs tends to interfere previously stored patterns (Lewandowsky, 1991; Lewandowsky & Li, 1995; McCloskey & Cohen, 1989). This interference is somewhat less strong in the VT model because there is a sparse coding in the network. This sparseness is a basic assumption of the model, and is crucial to account for the larger variance of old compared to new items strength distributions, which leads to a slope of the z-transformed ROC that is less than one. However, the aim of the VT model is to describe higher order representation in recognition memory, and we do not model the lower and perceptual levels of the cognitive system. We argue that a computational purpose of these earlier perceptual levels is to de-correlate similar inputs, so that the input to modeled layers are less influenced by catastrophic interference.

Even though no comparison is made with competing models in the current study, a note should be done regarding the characteristics of a more or less adequate model in reference to the VT. As described by Myung and Pitt (2004), certain criterions should be used to determine the strength of a specific model, of which goodness of fit, complexity and generalizability are conventionally stressed in the literature. The first should provide a measure that infers the models capability to make a fit of the underlying regularity of interest, and not be affected by noise (sampling error). Of course, since no comparison is made, such a measure is not included here (although, it was controlled whether the residuals were obtained within the 95% CI, which was the case in the majority of responses). However, the goodness of fit is related to the second criterion - complexity of the model. A complex model, that is, one with many parameters, tends to absorb random noise, and therein exhibit a higher fit without increasing the fit of the regularity of interest. Therefore, a simple model performs better in this respect. The variance theory uses 5 parameters (the standard deviation for high and low frequency, the mean net input to famous items, the standard deviation of input to the item layer and an activation threshold). The model was originally constructed to account for the mirror effect of recognition memory (Sikström, 2001), and was later applied to reaction time data (Sikström, 2004). As this is a founding principle in the model, and since it successfully accounted for data in both experiment 1 and 2 of the current study, it has proven to have predictive accuracy and to capture the processes that generated the data. Obviously, further research is necessary to establish the generalizability of the model in more detail. As the VT encompass the contribution of pre-experimental familiarity, future studies should comprise this dimension. Further, no study that directly assets the relation between fame and frequency and the respective layers of the variance theory has been conducted, and therefore should be carried out in future studies.

References

Achilles, E. M. (1920). Experimental studies in recall and recognition. *Archives of Psychology*, 27, 1-80.

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science, 2(6),* 396-408.

Bird, C. M., & Burgess, N. (2008). The hippocampus supports recognition memory for familiar words but not unfamiliar faces. *Curr Biol, 18*(24), 1932-1936.

Bird, C. M., Davies, R. A., Ward, J., & Burgess, N. (2011). Effects of pre-experimental knowledge on recognition memory. *Learn Mem, 18*(1), 11-14.

Bowles, B., Harlow, I. M., Meeking, M. M., & Kohler, S. (2012). Discriminating famous from fictional names based on lifetime experience: evidence in support of a signal-detection model based on finite mixture distributions. *J Exp Psychol Learn Mem Cogn, 38*(1), 78-91.

Clark, S. E., & Gronlund, S. D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin and Review, 3(1)*, 37_60.

DeCarlo, L. T. (2002). Signal detection theory with finite mixture distributions: theoretical developments with applications to recognition memory. *Psychol Rev, 109*(4), 710-721.

De Groot, A. D., Gobet, F., & Jongman, R. W. (1996). Perception and memory in chess: Studies inthe heuristics of the professional eye. Assen, The Netherlands: Van Gorcum & Co.

Dennis, S. (1995). The impact of the environment on the word frequency and the null list strength effects in recognition memory [Electronic version]. Noetica, 1. Retrieved 19 March 2007 from http://www2.psy.uq.edu.au/CogPsych/Noetica/Issue1/Environment.html

Diana, R. A., & Ranganath, C. (2011). Recollection, familiarity and memory strength: confusion about confounds. *Trends Cogn Sci, 15*(8), 337-338.

Diana, R. A., Reder, L. M., Arndt, J., & Park, H. (2006). Models of recognition: a review of arguments in favor of a dual-process account. *Psychon Bull Rev, 13*(1), 1-21.

Dunn, J. C. (2004). Remember-know: a matter of confidence. *Psychol Rev, 111*(2), 524-542.

Dunn, J. C. (2008). The dimensionality of the remember-know task: a state-trace analysis. *Psychol Rev, 115*(2), 426-446.

Egan, J. P. (1958). Recognition memory and the operating characteristic. *Technical note AFCRC-TN-58-51, Hearing and communication laboratory, Indiana University*.

Estes, W. K., & Maddox, W. T. (2002). On the processes underlying stimulus-familiarity effects in recognition of words and nonwords. *J Exp Psychol Learn Mem Cogn, 28*(6), 1003-1018.

Glanzer, M., & Adams, J. K. (1990a). The mirror effect in recognition memory: Data and theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(1), 5-16.

Glanzer, M., & Adams, J. K. (1990b). The mirror effect in recognition memory: data and theory. *J Exp Psychol Learn Mem Cogn, 16*(1), 5-16.

Glanzer, M., Kisok, K., & Adams, J. K. (1998). Response distributions as an explanation of the mirror effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(3), 633-644.

Glanzer, M., Hilford, A., & Kim, K. (2004). Six regularities of source recognition. *J Exp Psychol Learn Mem Cogn, 30*(6), 1176-1195.

Jang, Y., Mickes, L., & Wixted, J. T. (2012). Three tests and three corrections: comment on Koen and Yonelinas (2010). *J Exp Psychol Learn Mem Cogn, 38*(2), 513-523. doi: 10.1037/a0025880

Johnson, M. K. (2006). Memory and reality. *Am Psychol, 61*(8), 760-771.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychol Bull, 114*(1), 3-28.

Koen, J. D., & Yonelinas, A. P. (2010). Memory variability is due to the contribution of recollection and familiarity, not to encoding variability. *J Exp Psychol Learn Mem Cogn, 36*(6), 1536-1542.

Koen, J. D., & Yonelinas, A. P. (2013). Still no evidence for the encoding variability hypothesis: a reply to Jang, Mickes, and Wixted (2012) and Starns, Rotello, and Ratcliff (2012). *J Exp Psychol Learn Mem Cogn, 39*(1), 304-312.

Lewandowsky, S. (1991). Gradual unlearning and catastrophic interference: A comparison of distributed architectures. In W. E. Hockley & S. Lewandowsky (Eds.), *Relating theory and data: Essays on human memory in honor of Bennet B. Murdock* (pp. 445-476). NJ: Hillsdale.

Lewandowsky, S. , & Li, Shu. (1995). Catastrophic interference in neural networks: Causes, solutions and data. In F. N. Dempster & C. J. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 329-361). San Diego: Academic press.

Mandler, G. (1980). Recognizing: The Judgment of previous occurrence. *Psychol Rev, 87*, 252-271.

Mandler, G. (2008). Familiarity breeds attempts: A critical review of dual process theories of recognition. *Perspectives in Psychological Science 3 (5)*, 392–401

McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning paradigm. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation* (Vol. 24, pp. 109-164). San Diego: Academic Press.

Mickes, L., Wais, P. E., & Wixted, J. T. (2009). Recollection is a continuous process: implications for dual-process theories of recognition memory. *Psychol Sci, 20*(4), 509-515.

Mitchell, K. J., & Johnson, M. K. (2000). Source monitoring: Attributing mental experiences. In E. Tulving & F.I.M. Craik (Eds.), *The Oxford Handbook of Memory* (pp. 179-195). New York: Oxford University Press.

Onyper, S. V., Zhang, Y. X., & Howard, M. W. (2010). Some-or-none recollection: Evidence from item and source memory. *J Exp Psychol Gen, 139*(2), 341-364.

Parr, W. V., Heatherbell, D., & White, K. G. (2002). Demystifying wine expertise: olfactory threshold, perceptual skill and semantic memory in expert and novice wine judges. *Chem Senses, 27*(8), 747-755.

Pratte, M. S., & Rouder, J. N. (2012). Assessing the dissociability of recollection and familiarity in recognition memory. *J Exp Psychol Learn Mem Cogn, 38*(6), 1591-1607.

Qin, J., Raye, C. L., Johnson, M. K., & Mitchell, K. J. (2001). Source ROCs are (typically) curvilinear: comment on Yonelinas (1999). *J Exp Psychol Learn Mem Cogn, 27*(4), 1110-1115.

Ratcliff, R., Van Zandt, T., & McKoon, G. (1995). Process dissociation, single-process theories, and recognition memory. *J Exp Psychol Gen, 124*(4), 352-374.

Reder, L. M., Nhouyvanisvong, A., Schunn, C. D., Ayers, M. S., Angstadt, P., & Hiraki, K. (2000). A mechanistic account of the mirror effect for word frequency: A computational model of remember-know judgments in a continuous recognition paradigm. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 26*, 294-320.

Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends Cogn Sci, 11*(6), 251-257.

Rugg, M. D., & Yonelinas, A. P. (2003). Human recognition memory: a cognitive neuroscience perspective. *Trends Cogn Sci, 7*(7), 313-319.

Sikström, S. (2001). The variance theory of the mirror effect in recognition memory. *Psychon Bull Rev, 8*(3), 408-438.

Sikström, S. (2004). The variance reaction time model. *Cogn Psychol, 48*(4), 371-421.

Skinner, E. I., & Fernandes, M. A. (2007). Neural correlates of recollection and familiarity: a review of neuroimaging and patient data. *Neuropsychologia, 45*(10), 2163-2179.

Slotnick, S. D. (2010). "Remember" source memory ROCs indicate recollection is a continuous process. *Memory, 18*(1), 27-39.

Slotnick, S. D., Klein, S. A., Dodson, C. S., & Shimamura, A. P. (2000). An analysis of signal detection and threshold models of source memory. *J Exp Psychol Learn Mem Cogn, 26*(6), 1499-1517.

Smith, D. G., & Duncan, M. J. (2004). Testing theories of recognition memory by predicting performance across paradigms. *J Exp Psychol Learn Mem Cogn, 30*(3), 615-625.

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

Squire, L. R., Wixted, J. T., & Clark, R. E. (2007). Recognition memory and the medial temporal lobe: a new perspective. *Nat Rev Neurosci, 8*(11), 872-883.

Starns, J. J., Rotello, C. M., & Ratcliff, R. (2012). Mixing strong and weak targets provides no evidence against the unequal-variance explanation of ZROC slope: a comment on Koen and Yonelinas (2010). *J Exp Psychol Learn Mem Cogn, 38*(3), 793-801.

Stenberg, G., Hellman, J., & Johansson, M. (2008). The memorability of names and the divergent effects of prior experience. *The European Journal of Cognitive Psychology, 20* (2), 312-345.

Stenberg, G., Hellman, J., Johansson, M., & Rosen, I. (2009). Familiarity or conceptual priming: event-related potentials in name recognition. *J Cogn Neurosci, 21*(3), 447-460.

Tulving, E. , & Thompson, D. M. (1973). Encoding specificity and retrieval process in episodic process. *J Exp Psychol, 87*, 353-373.

Valentin, D. & Chollet, S. (2007). Expertise and memory for beers and beer olfactory compounds. *Food Quality and Preference, 18(5)*, 776-785.

Wais, P. E., Mickes, L., & Wixted, J. T. (2008). Remember/know judgments probe degrees of recollection. *J Cogn Neurosci, 20*(3), 400-405.

Westmacott, R., & Moscovitch, M. (2003). The contribution of autobiographical significance to semantic memory. *Mem Cognit, 31*(5), 761-774.

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychol Rev, 114*(1), 152-176.

Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: evidence for a dual-process model. *J Exp Psychol Learn Mem Cogn, 20*(6), 1341-1354.

Yonelinas, A. P. (2001). Components of episodic memory: the contribution of recollection and familiarity. *Philos Trans R Soc Lond B Biol Sci, 356*(1413), 1363-1374.

Yonelinas, A. P., Aly, M., Wang, W. C., & Koen, J. D. (2010). Recollection and familiarity: examining controversial assumptions and new directions. *Hippocampus, 20*(11), 1178-1194.

Yonelinas, A. P., & Parks, C. M. (2007). Receiver operating characteristics (ROCs) in recognition memory: a review. *Psychol Bull, 133*(5), 800-832.

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.

Notes

[1 & 2] In the current experiment there is a reliable change in the criterion, however, the computational model assumes that movement of the distributions reflects this shift.

Paper IV

The Multidimensional Signal Detection Theory

Johan Hellman[1], Sverker Sikström[1]
[1]Lund University, Sweden

Signal detection theory (SDT) and the Dual Process SDT (Yonelinas, 1994) are the most influential theoretical frameworks for quantifying the underlying familiarity distributions. However, neither provides a detailed account for the basic finding that the old item distributions have larger variability than that of the new items, a phenomenon that has been accounted for by the hypothesis of encoding variability (Wixted, 2007) or by assuming dual processes in recognition memory (Yonelinas, 1994). We present the Multidimensional Signal Detection Theory (the MSDT) that suggests that the underlying familiarity distributions can be described by a binomial density function, rather than a normal density function that is commonly assumed. Attention is modeled by focusing the signal to a few of the dimensions. The model accounts for performance, item variability (ROC curves) and response variability in attentive and inattentive persons and suggests a positive relation between attention, distinct neural representation, memory performance and dopamine whereas attention is negatively related to new/old response variability and overall response variability. We tested these predictions on attentive and inattentive participants, and found a lower z-slope in attentive people.

Acknowledgements
This research was supported by the Swedish Research Council (grant to project "Semantic Spaces as a Method in Psychology") and the Wallenberg foundation (grant to project "Noise as a Tool to Improves school performance in Inattentive Schoolchildren").

In the literature on recognition memory, several theories have provided accounts of encoding and retrieval processes and item variability. However, such models solely rely on processes involved in item familiarity, may they be based on a single variable of memory strength (signal detection theory) or dual processes in recognition memory (for review, see Yonelinas, 2002; Yonelinas & Parks, 2007). In the present paper, we will review models of recognition memory, describe their accounts of variability in item memory decisions, elucidate their limitations and introduce a new model called the Multidimensional Signal Detection Theory (the MSDT). An important contribution of the MSDT is that it modulates attention and how it interacts with item recognition, as well as item variability and response variability in people with normal and impaired attention (i.e., ADHD). Thus, besides providing a plausible and extended account of item variability, the MSDT introduces a possible integration of two usually separated fields: the study of bias and sensitivity in recognition memory and the understanding of ADHD symptomology.

Signal detection theory and dual-process theory

Signal detection theory (SDT) has been the prominent theoretical framework for understanding recognition memory decisions. In its conventional form it states that memory can be described by two equal-variance Gaussian distributions, reflecting two random variables of signal and noise (old and new items) along a familiarity continuum; the equal-variance SDT (or the EVSD model). Items that exceed a decision criterion, which denote a certain level of familiarity, are accepted as old, otherwise they are rejected as new. However, empirical work has demonstrated that the familiarity distributions require an unequal-variance model (the UVSD model), because the variance of the old items exceeds that of the new items (Mickes, Wixted, & Wais, 2007; Ratcliff, Sheu, & Gronlund, 1992; Yonelinas & Parks, 2007).

Item variability is commonly asserted with the receiver-operating characteristics (ROC) approach. An ROC curve relates the hit rates to the false alarm rates over different confidence intervals. If the hit and false alarm rates are z-transformed, the ROC equals an approximately straight line with a slope (z-slope) at or below 1.0. Equal variability of the new and old item distributions results in a z-slope of 1.0, whereas a higher variability in the old item distribution generates a z-slope below 1.0 (Egan, 1958). Although higher z-ROC slopes have been observed in other domains (Swets, 1986), the most frequent observations in memory research are z-slopes below 1.0.

To account for differences in item variability, the verbally formulated account of encoding variability (the encoding variability hypothesis, here on after abbreviated EVH) has been suggested (Wixted, 2007). That is, as the memory strength for some items increase more than other items during study, the old item distribution gets more varied than that of the new items. But, the current version of SDT does not provide a detailed theoretical understanding for why the old distribution has a larger variability than the new (Koen & Yonelinas, 2010, 2013), and the EVH only asserts that the difference in variability is related to differences in item familiarity, and not how this difference is mediated.

A more recent account for item variability is the Dual Process Signal Detection theory (Yonelinas, 2001). It describes recognition memory as consisting of two memory processes: familiarity and recollection. Familiarity is described as a graded process by which an assessment of memory strength for the presented item determines whether the item will be endorsed as old or rejected as new. Familiarity results in both correct and incorrect responses with varying degrees of confidence, and is described by a signal detection process. Recollection, on the other hand, enables retrieval of additional contextual information about the study event and results in a higher proportion of high confident hits and no errors. Recollection is described as a threshold process and items that do not reach the recollective threshold are recognized by familiarity. The two processes have different effects on performance. Familiarity leads to both correct and incorrect responses whereas recollection elevates accuracy, which results in an increased old item variability. Thus, the DPSD does not need to rely on the EVH. However, the DPSD must assume that recollection is a threshold process to vindicate the observed differences in old and new item variability, and this description has been debated (Klauer & Kellen, 2010; Malmberg, 2002; Mickes, Johnson, & Wixted, 2010; Mickes, Wais, & Wixted, 2009; Slotnick, 2010; Slotnick, Klein, Dodson, & Shimamura, 2000; Yonelinas, Dobbins, Szymanski, Dhaliwal, & King, 1996; Yonelinas & Parks, 2007).

It should be noted initially that we have no pretention on the debate regarding whether recognition memory constitute one or two retrieval processes, or how the UVSD and the DPSD differ in their accounts thereof. The primary reason for comparing the MSDT with these models is that they are seemingly the most influential models of item variability. We argue that the UVSD is limited in its account of ROC data because it must rely on an auxiliary hypothesis (the EVH will be further discussed below), and the DPSD can be questioned because it assumes a threshold process to account for empirical data. We do, however, make no claims on whether the threshold process assumption is correct or not, rather we view it as an additional parameter which may be redundant for the explanation of ROC data. Further, the EVH has rarely been tested, and when put under empirical test, it has been questioned (Koen & Yonelinas, 2010; 2013). We present a new multidimensional extension of SDT called the multidimensional signal detection theory (the MSDT). The MSDT provides an elaborated account of ROC data and response variability with novel predictions, as will be described below, by integrating conventions of SDT models with attentional functioning and specificity of neural representations of memory. The MSDT suggests synergy effects in SDT and a view of recognition memory as multidimensional, and as based on a binomial distribution. As will be described further ahead, the MSDT generates predictions for ROC data different from those derived by SDT and the DPSD.

The MSDT

Below the MSDT will be described and we will show that the model can explain performance and item variability (i.e., ROC data – variability in the new versus the old item distribution), as well as response variability and how these findings interact with attentiveness. The MSDT sets up some new predictions for both item- and response variability, as will be delineated further ahead.

In the MSDT, information is distributed represented in *N* number of nodes (denoted with the subscript i), where each node represents a feature in the stimuli. The majority of the nodes receive noise input only, whereas dependent on attention, a subset of consisting of (*a*) number of nodes has an input of both signal and noise for old items. The summed signal over all nodes is (*S*), which is evenly distributed over the *a* number of nodes $s_{1,2,\ldots a} = S/a$. All other nodes do not receive a signal: $s_{a+1,a+2,\ldots N} = 0$. For new items, no node receives a signal. The output of each node is the activity, where a node is active if the input exceeds the activation threshold (*t*), and otherwise it is inactive. Thus, each node is either active or inactive, which is implemented by a stepwise non-linear activation function. An old response is made if one or more nodes are active, where the ROC-curves can be generated / plotted by changing the activation threshold. Attention is modulated by the number of nodes that receives a signal, denoted the attention parameter (*a*). The total amount of signal (*S*) is equivalent for inattentive and attentive people. However, in attentive people the signal is focused to fewer nodes that receive a larger signal, and in inattentive people the signal is spread to more nodes where each node receives a smaller signal. The noise is normally distributed with a mean of zero and variance of one. In the MSDT, the probability function for a hit is (Eq 1):

$$P(\text{"Old"}|\text{Old}) = \left[\sum_{i=1}^{N}(s_i + n_{i)}) > t\right] > 0$$

New responses can similarly be calculated with this function given that $S = 0$.

The distribution of active nodes is described by a binomial distribution rather than a normal distribution as is commonly assumed in SDT. We view this as an important and interesting aspect of the model because it provides a simple and elegant understanding of why the variability of the old distribution is larger than the new distribution, due to the fact that the binomial distribution is positively skewed, which is particularly salient when the number of active nodes is small.

The MSDT: discussion of parameters

The MSDT is described by three parameters: the activation threshold (*t*), the attention parameter (*a*), and the signal (*S*). In addition, the number of nodes (*N*) is a scaling variable. These parameters can be compared to the threshold (C), the old item variability (hereby after denoted σ_O) and performance (d') in SDT, although they have different theoretical foundations.

Attentional skill is modulated by the attention parameter (*a*), where inattentive people (e.g., corresponding to attentional deficit, such as ADHD, but also people with low attentional ability) exhibit a higher number of nodes receiving a signal compared to attentive persons. For attentive people (low values of *a*), the signal is focused to fewer nodes and results in stronger signal to these few nodes, which makes it more likely that at least one node receives a signal plus noise that exceeds the activation threshold. Thus, attentive people will have a distinct neural representation of the presented item because the signal is focused to a small number of nodes whereas inattentive people have a more spread span of signal in the distribution of nodes, so the signal to a single node is smaller, inducing a less specific neural representation of the test item. The MSDT suggests that the assumption of differences in signal distribution can be related to dopamine levels, where attentive people have higher levels of dopamine which results in a higher input-output gain leading to a shaper neural representation (Servan-Schreiber, Printz, & Cohen, 1990; Servan-Schrieber, Bruno, Carter, & Cohen, 1998; Li & Sikström, 2002). This is elaborated further ahead.

The activation threshold (*t*) reflects the response criterion as well as the threshold for node activity. A node becomes active given that the input consisting of signal plus noise (or noise only) exceeds the activation threshold and an item receives and old response when at least one node is active (independent on whether the node is activated by noise only or signal plus noise), whereas a new response is given if no node is active.

For an unbiased response to occur, that is neither to conservative nor to liberal, the activation threshold needs to be appropriately placed. This placement depends on the other parameters, particularly the attention parameter, of the model in a non-linear way, and finding an unbiased placement threshold requires a numerical solution of the equation describing the model. We introduce an attention related bias term on the activation threshold ($t_{unbiased}(a, S, N)$), which is placed so that the predicted values fulfills the following criteria p(FA) = 1-p(H), i.e., the threshold where hit rates and false alarm rates are symmetrical around 50%. The bias term is calculated by numerically solving the model parameters given the constraints mentioned above, and the value of this term depends on the other parameters. The threshold equals the sum of the bias term and an unbiased term, where the former in isolation yields a mirror effect (i.e., when $t_{unbiased}$=0) and the latter produces a conservative response when it is negative ($t_{unbiased}$<0) and liberal response when it is positive ($t_{unbiased}$>0) (Eq. 2):

$$t = t_{unbiased} + t_{biase}$$

The threshold (*t*) is affected by (*a*), so that an increase in number of nodes receiving a signal decreases (*t*). Thus, inattentive people with a less focused signal have a lower activation threshold, resulting in a higher number of active noise-only nodes (i.e., false alarms). This relates to the variability of the latent distributions, meaning that an increase in (*a*) results in lower variability in the old item distribution.

That inattentive people exhibit more false alarms than attentive people is not predicted by the UVSD or the DPSD, however, there is support for the hypothesis in the literature. Previous studies have observed higher false alarms in ADHD as compared to healthy controls, albeit primarily on children and in the auditory domain (Breier, Gray, Klaas, Fletcher, & Foorman, 2002; Gray, Breier, Foorman, & Fletcher, 2002; Uebel et al., 2010). Breier and colleagues (et al., 2002) tested children with ADHD and healthy controls on auditory recognition where a central masking manipulation was applied and compared to a non-masking condition. For children with ADHD, false alarms increased in comparison with the controls in the central masking condition, but not in the no-masking condition. Further, using a Go/Nogo task on children with ADHD, non-affected siblings and healthy controls, Uebel (et al., 2010) observed reliably higher false alarms, longer response time and higher response variability in children with ADHD as compared to non-affected siblings and healthy controls. In the MSDT, the difference in response variability is modeled by introducing variability in the activation threshold (as is elaborated further ahead).

Strength variables; such as study time, how easily the material is to encode, etc. depends on the signal (*S*). Thus, (*S*) is higher for repetition (high familiar as compared to less familiar items).

The number of nodes (*N)* is a scaling variable, because for large values it does not make significant influence on the fit of the model, however, it scales other parameters in the model. Table 1 shows how well the model fit (here reported as mean-square error) a dataset (see below for a detailed description of this dataset) depending on a large range of values of (*N*), and the associated fitted values of (*a*), (*t*) and (*S*). The fit of the model (*mse*) is in essence unaffected by (*N*) whereas the other variables (*a*, *t*, and *S*) show small increases as *N* increase over magnitudes.

Table 1. Relation between (N) and the model parameters

(N)	mse	a(Inatt)	a(Att)	(S)	(t)
500	0.0032	2.498	2.003	7.789	1.546
1000	0.0032	2.638	2.141	8.517	1.689
2500	0.0033	2.733	2.262	9.343	1.842
10000	0.0034	3.063	2.586	11.073	2.093

Note. mse denotes the mean-square error, a(Inatt) and a(Att) the number of nodes receiving a signal for inattentive and attentive respectively, (S) the signal, and (t) the activation threshold.

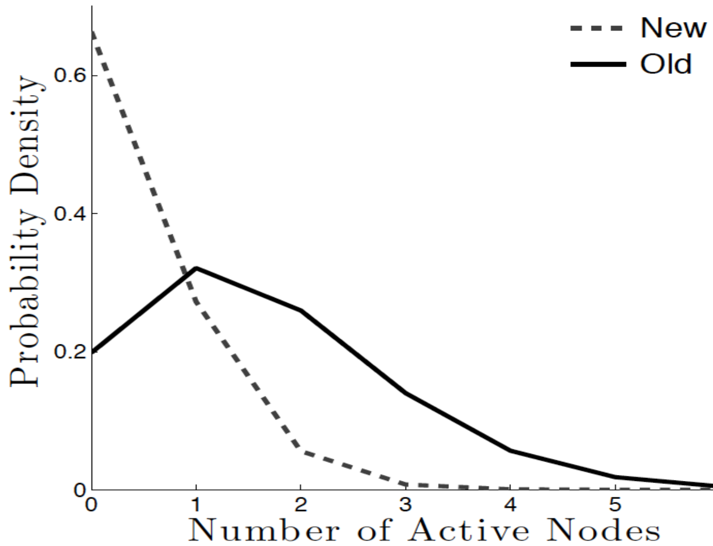
Model description and predictions

A core feature in the MSDT is that the probability that nodes are active for old items is larger than the probability that they are active for new items, which leads to an increase in the old as compared to the new item variability. When the probability for nodes to be active is zero, then the variability is also zero. This variability increases as the probability for the nodes to be active increases, and reaches a peak at 50%. However, consistent with our understanding of the neural system (Quiroga, Kreiman, Koch, & Fried, 2008; Quiroga, Reddy, Kreiman, Koch, & Fried, 2005) we assume a sparse representation with a low ratio of active nodes. Figure 1 plots the probability density function of the number of active nodes at retrieval. New items have substantially lower variability in node activity with a maximal probability at null, and old items have a peak at a larger number of active nodes at retrieval.

Performance in attentive and inattentive

The panels in Figure 2 show how the distribution of signal changes with an increase in (*a*), where the magnitude of the signal (for nodes receiving a signal) decreases as (*a*) increases. The noise (in gray) and the signal input (in red) are plotted in relation to (*t)* (the solid line) on the y-axis for a number of nodes along the x-axis. This is displayed for attentive (Panel A), inattentive (Panel B) and highly inattentive persons (Panel C).

Figure 1. Variability of active nodes

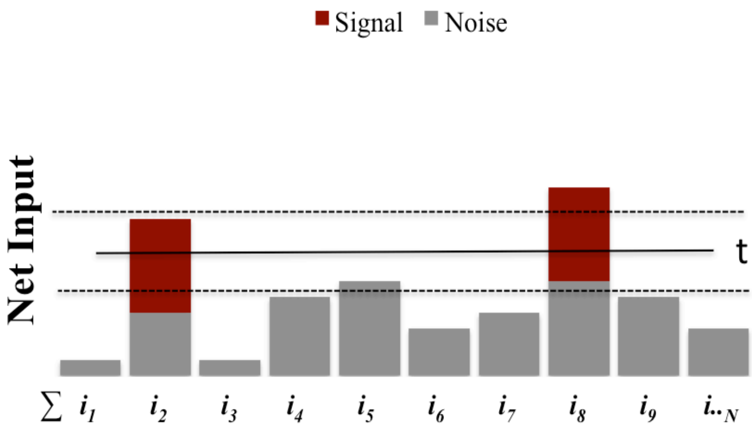


Note. Solid black line represents old items and dotted grey line new items.

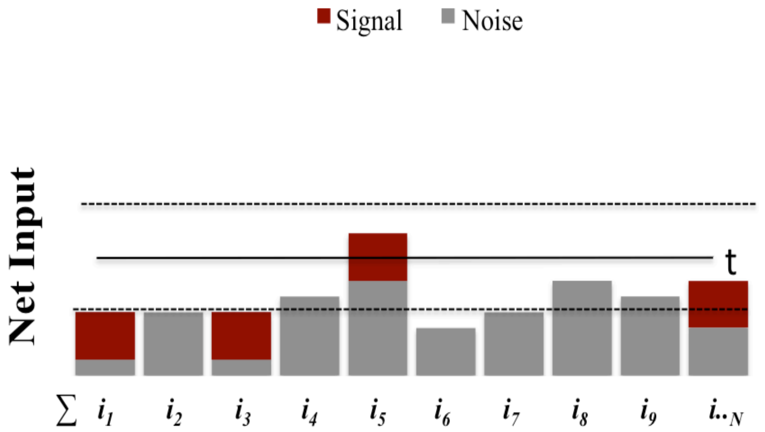
The figure visualizes how signal plus noise is allocated depending on the degree of attentiveness. The area within the dotted lines shows a hypothetical variability in the activation threshold.

Figure 2. Net input to nodes over attentiveness

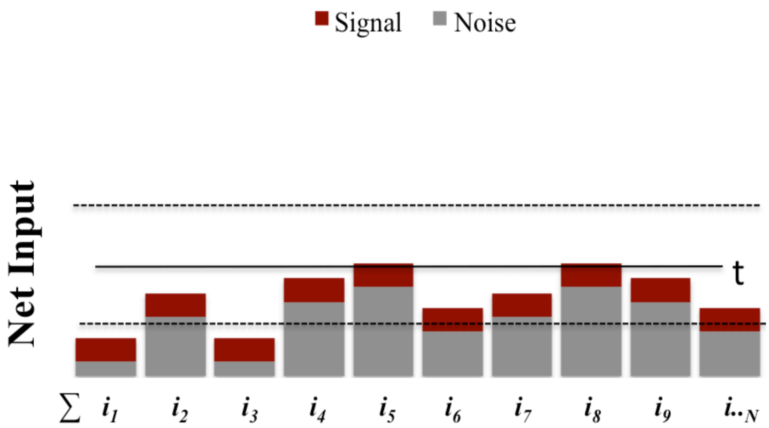
Panel A. Attentive



Panel B. Inattentive



Panel C. Highly Inattentive



Note. In all panels, grey stables represent noise input and red stables signal input. The horizontal solid line represents the activation threshold (t), and the two dotted lines represent the area of variability in (t). Below the x-axis the index for each node is displayed, from 1 to N. Activity in all nodes exceeding (t) are summed.

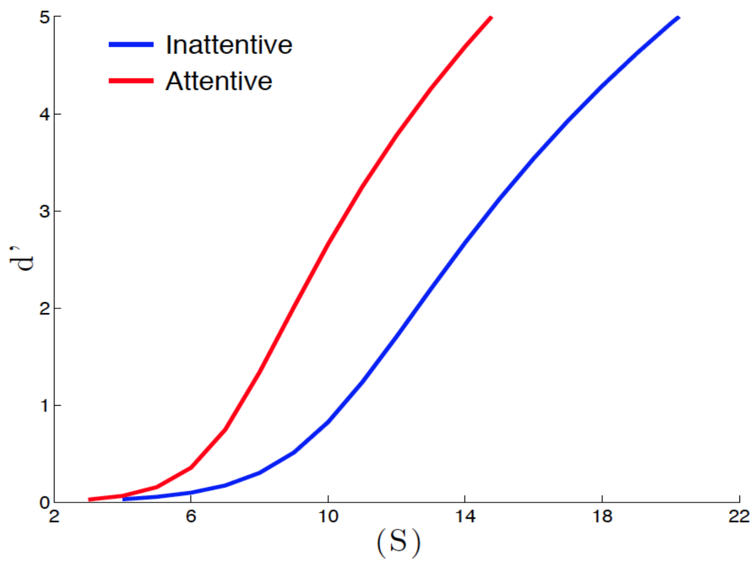
The noise and the total signal is the same for all levels of attention, although the distribution of signal depends on (*a*). For high attention (Panel A), few nodes are activated as the result of low values of (*a*) of which a relatively large proportion constitute signal. The focused allocation of signal, that occurs for low values of (*a*), results in a positively skewed distribution of node activity, which can be described by a binomial distribution.

When the number of active nodes increases (see Panel B) performance variability increases, which also results in a distribution that approximates a Gaussian distribution of node activity (as is assumed in SDT). Panel C plots signal and noise allocation for highly attentive people, and serves an informative example. Here, all nodes consist of both signal and noise due to high values of (*a*), leading to poor representation of the test item, and therefore, lower memory performance as compared to lower values of (*a*). This also highlights the difference between the MSDT and conventional SDT, as the latter is analogous to the SDT for $N = 1$ and $a = 1$. Thus, the MSDT provides a detailed account of recognition memory and ROC data where (*a)* is related to the distribution of the signal, making it a multidimensional extension to the conventional SDT.

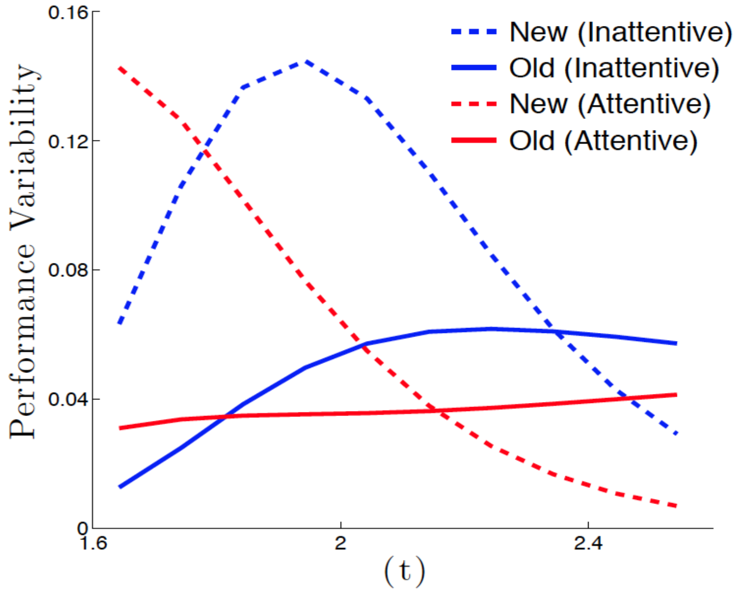
The MSDT predicts higher memory performance in attentive than inattentive people. Figure 3, Panel A plots recognition performance (d') as a function of signal (*S*) and the figure shows a higher accuracy for attentive as compared to inattentive persons. For the latter group, the signal is distributed over a larger number of nodes and therefore reduces the probability that at least one node is activated, which leads to a decline in performance. This is consistent with the finding that sparse neural network provides higher performance than densely represented networks (Hertz, Krogh, & Palmer, 1991).

Figure 3. Performance and response variability

Panel A. Performance and (*S*)



Panel B. Performance variability in attentive and inattentive



Note. In both panels, red lines represent attentive and blue line inattentive. Panel B, dotted lines represent new items and solid lines old items.

An extended account of variability

Variability is a key concept for the understanding of memory, however, it is commonly investigated along different paths of enquire. On the one hand, item variability is studied by the use of Receiver-operating characteristics (ROCs), where focus is put on how variability of responses to new and old items differ. On the other hand, differences in response variability, that is, how yes and no responses to test items differ over subjects for several different tasks (recognition performance, reaction time, etc.), are commonly separated from the study of item variability. The MSDT provides a unified account of both response- and item variability, as is described below.

Response variability

Increased response variability in ADHD patients is a common finding in the literature (Castellanos et al., 2005; Leth-Steensen, Elbaz, & Douglas, 2000). This has been observed both between and within participants in children (Williams et al., 2005) and adults (Hultsch, MacDonald, & Dixon, 2002). This effect is also observed over a wide span of tasks. For instance, Leth-Steensen and colleagues (Leth-Steensen et al., 2000) reported large response time variability and slightly larger mean response times in ADHD as compared to controls in a color-key matching task. Further, ADHD diagnosed participants have been shown to exhibit larger response variability than controls in the Eriksen flanker task (Castellanos et al., 2005), in finger-tapping and anticipation tasks (Toplak & Tannock, 2005), and in a time perception task in children (Toplak et al., 2003). Also, the population exhibits higher than normal variability in attention-demanding tasks, such as sustained and selected attention in continuous performance and conjunctive search tasks, as well as orienting and executive attention in cost-benefit and Stroop tasks, respectively (Tsal, Shalev, & Mevorach, 2005).

The MSDT predicts higher response variability for inattentive than attentive people due to an increased variability in the activation threshold, which results in a higher variability in correct and incorrect recognition responses. This occurs because inattentive have a signal span that is allocated over a larger number of nodes than attentive, leading to a greater change in the number of active nodes due to the changes in the placement of the activation threshold. Thus, we model response variability by introducing variability in the activation threshold.

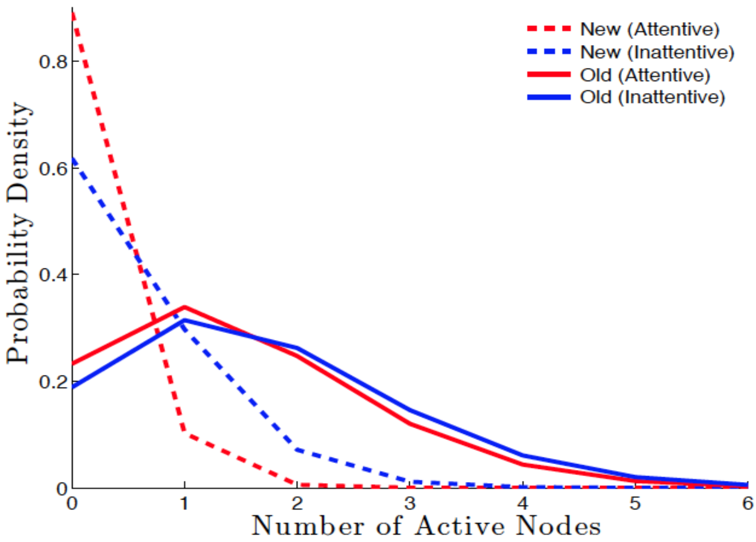
A novel and important prediction in the MSDT is that inattentive people have larger response variability for new than for old items. This is predicted because new items have lower variability, which makes it relatively more influenced by a change in the activation threshold as compared to old items. Response variability according to the MSDT is visualized in Figure 3, Panel B, and shows a change in performance following a change in the activation threshold ($\Delta t = 0.1$). This constitutes a measure of variability for old and new items for attentive and inattentive persons for different values of (t). For all values of the activation threshold, the variability in performance is larger for inattentive than attentive people. For low values of (t) (i.e., liberal judgments), new item variability is higher than that of old items and this difference is larger in inattentive than attentive persons. For moderate values of (t), old item variability increases for attentive whereas the opposite happens for inattentive. For high values of (t) (i.e., conservative judgments), old item performance variability exceeds that of the new items, where inattentive exhibit higher values than attentive persons. Thus, the MSDT suggests that the increase in response variability in inattentive to a larger extent is driven by changes in the false alarm rates compared to changes in hit rates. In contrast, response variability in attentive is to a larger extent modulated by hit rates relative false alarm rates. These differences, then, occur because there is a larger variability in (t) for inattentive than attentive people and the model specifically predict that inattentive people implements a more liberal response criterion than attentive people.

Item variability

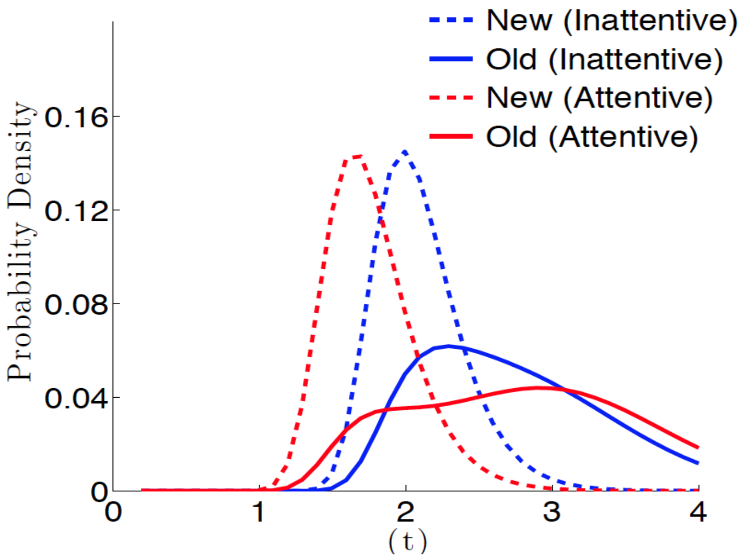
Here we show how variability of the familiarity distributions for attentive/inattentive persons, and new/old responses (i.e., the z-slope) depends on the attentiveness parameter (*a*). The slope of the z-ROC curves depends on the number of nodes receiving signal where new items have a higher variability in familiarity than old items. Figure 4, Panel A, plots the probability density of number of activated nodes, divided into attentive and inattentive people. As can be seen, one active node has the largest probability of occurring for old items, and this function is similar for both high and low levels of attention. For new items, on the other hand, the largest probability occurs when no nodes are active. For attentive people this probability decreases sharply when the number of nodes grows, whereas for inattentive people the slope gets shallower for an increase in nodes. Thus, differences in new to old item variability in attentive and inattentive are primarily a result of changes in the new than the old item distribution.

Figure 4. New and old item variability

Panel A. Probability density of number of active nodes



Panel B. Probability density of familiarity over the activation threshold



Note. Panel A shows the probability density over the number of active nodes and Panel B shows the probability density of familiarity over (t). Solid lines represents old items and dotted lines new items. Attentive in red and inattentive in blue color.

Thus, the MSDT predicts higher z-ROC slopes for inattentive persons because there is a smaller difference in new and old item variability than for attentive people, i.e., the relatively small change in the new item variability influences the z-ROC slope. Accordingly, Figure 4, Panel B plots the probability density of familiarity as a function of (*t*). This figure reveals that inattentive people are more inclined to endorse a higher number of new items as old compared to attentive people, as described above.

Describing recognition memory with a binomial distribution

The literature relating to variability of new and old items typically assumes that the underlying distribution is Gaussian (Yonelinas & Parks, 2007). This assumption is based on two arguments. First, the underlying distribution is normal because the summation of arbitrary distributions generates a Gaussian distribution given that the summation is done over a sufficient number of nodes. Second, a normal distribution predicts that z-transformed ROC curves are linear, which is consistent with empirical data (but non-linear z-shapes has been observed, see Yonelinas & Parks, 2007).

The MSDT suggests that nodes are either active or not, which implies a binomial underlying distribution. This distribution has a precise mathematical distribution, where data near the limits of probability ($0 \leq probability \leq 1$) generates distributions that have a lower variability than data in the middle of the distribution. The MSDT assumes that the probabilities of active nodes are near the lower limit, based on the argument that the representation of a word in the brain is relatively sparse (Rolls & Treves, 2011), and that the probability of actives nodes for old items is larger than for new items. The fact that MSDT predicts a binomial distribution has several important implications, as described below.

The slope of the z-ROC curve does not equal to ratio of the old and new standard deviation. The z-slope is typically interpreted as a measure of the ratio of the new to old standard deviation (Yonelinas & Parks, 2007), but the MSDT implies that this is inaccurate. We compared two different ratios related to item variability by fitting the parameters so that the z-slope = 0.8, d' = 2 and a mirror effect occurs (i.e., hits and false alarms are symmetrical around 50%; FA = 1-H). The result is presented in Table 2. As is evident from the table, the z-slope overestimates the ratio of new to old standard deviation, so that a z-ROC slope of 0.8 corresponds to a ratio of new to old standard deviations of approximately 0.48 and a ratio of the active new to the active old node probability of 0.31.

Table 2. Two different ratios of item variability

Measure	Value
z-ROC slope	0.80
StD(N)/StD(O)	0.48
$pN/pO^2=pN/pO^2$	0.31

Note. Values are derived by setting d' = 2, z-ROC slope = 0.8 and FA = 1-H. Std(N) and StD(O) denotes the standard deviation of new and old items and $pN/pO^2=pN/pO^2$ is the ratio of the active new to the active old node probability.

The ratio of new to old number of active nodes is related to the ratio of new and old item variability. Given a binomial distribution, the ratio of the new to old standard deviation is approximately 0.5 when d' = 2, the z-slope equals 0.8, that the hits and false alarm rate are made symmetrical around 50% (i.e., a mirror effect is implemented where FA = 1 - H), and a large number of nodes. For these parameter settings, the ratio of the active new to the active old node probability is 0.31.

To describe the relation between net input and active nodes, it is possible to approximate the binomial distribution with a normal distribution with a variance of (Eq. 3):

$$v_s = p(1 - p)N$$

This approximation is reasonable given that p (1 – p) $N > 5$ (i.e., the number of active nodes should be at least five, or preferably, ten). The ratio of the new (subscript n) to old item (subscript o) variability can then be expressed as (Eq. 4):

$$\frac{v_{sn}}{v_{so}} = \frac{p_n(1 - p_n)}{p_o(1 - p_o)}$$

Given a low probability that the nodes are active (e.g., $p < .1$), this expression can approximately be simplified to (Eq. 5):

$$\frac{v_{sn}}{v_{so}} = \frac{p_n}{p_o}$$

This expression suggests a simple interpretation of the ratio of the new and old familiarity variability, namely that this constitutes a measure of the relative number of activated nodes (*a*) in the new and old distributions. However, in a network with low number of active nodes this derivation turns out to be incorrect (see Figure 4, Panel A, showing that the number of active nodes is lower than 5). Thus, the relative number of active nodes (that is manipulated with *t*) overlaps with the new and old item familiarity, and this value differs from the slope of the z-ROC. Thus, using a normal distribution to approximate a binomial distribution is inadequate given the parameter settings described above.

The z-transformed ROC curve is approximately linear. The empirical fact that the z-ROC curve is approximately linear in shape has been used as an argument that the underlying distribution is normal (Lockhart & Murdock, 1970; Yonelinas & Parks, 2007). Using the parameter setting described above, we notice that the z-ROC curve proximate linearity to such a degree that it is practically impossible to distinguish from a linear function given the resolution of current empirical data. The (linear) correlation between the z-transformed hit rates and z-transformed false alarm rates for different confidence criteria are 0.994 for attentive and 0.997 for inattentive people. Thus, a linear z-shape cannot efficiently be used to distinguish between a normal and binomial distribution with low variability (e.g., *a*) independent of *N*, suggesting that a linear z-ROC shape does not necessarily reflect an underlying Gaussian distribution. For a discussion on whether parametric models of memory strength are empirically testable, see Rouder, Pratte and Morey, (2010), and Wixted & Mickes (2010).

Interpreting unequal variance data as non-linear activation functions. The MSDT suggests that the larger variability in familiarity for old than new items occurs because the nodes have a non-linear activation functions, where the variability of old items exceeds that of new items due to low values of (*a*). This is consistent with empirical data where the old item distribution has larger variability than that of the new items (Hilford, Glanzer, Kim, & DeCarlo, 2002; Yonelinas & Parks, 2007). However, the MSDT can account for both equal and unequal variability in the familiarity distributions. Figure 5, Panel A, displays z-ROC curves for attentive and inattentive persons, demonstrating that differences in activity of the underlying nodes and the input of signal and noise affects item variability differently for high and low levels of attention.

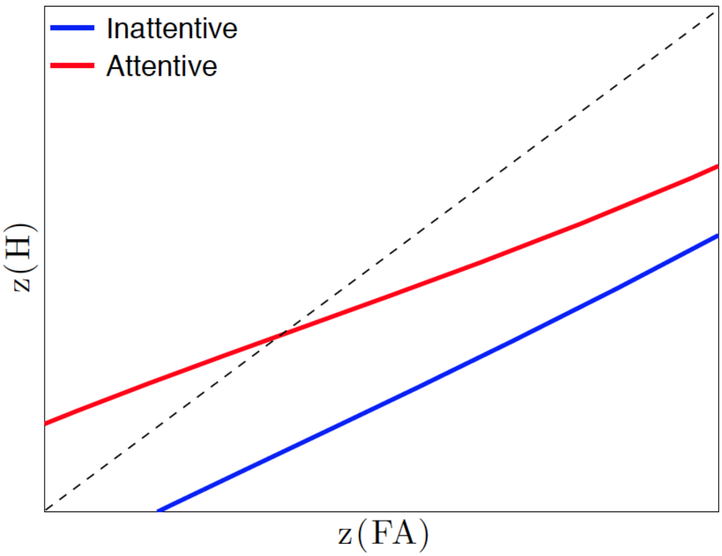
A specific relation between the z-slope and performance. There has been a debate on how the z-slope is related to performance (Glanzer, Kim, Hilford, & Adams, 1999; Ratcliff, McKoon, & Tindall, 1994), which will be described further ahead (see 1.4.1). However, it should be mentioned here that the MSDT suggests a specific relation between the z-slope and performance. The slope of the z-ROC is plotted as a function of d' in Figure 5, Panel B. We investigate this empirical value of the z-slope at 0.8 by constraining the values of (*a)* and (*N* = 80000) so that the z-ROC slope = 0.8 when d' = 2. According to the MSDT, the z-ROC slope decreases rather steeply for 0 < d' < 0.5 and gets smoother in the area of 0.5 < d' < 2, but this decrease becomes shallower for higher d' values, up to d' = 3.

Empirical values of z-ROC slope indicate sparse node activation. Several studies of item variability have concluded that item z-ROC slopes approximates a value of 0.8 (Dunn, 2004; Mickes et al., 2007; Rotello, Macmillan, & Reeder, 2004). Therefore, a formal model of item variability must account for z-ROC values neighboring this value. Figure 5, Panel C plots z-ROC slope as a function of the number of nodes receiving a signal (*a*), for d' = 2, and (*a*) varying from 0 to 10. The figure shows that a change in (*a*) strongly affects the z-ROC slope, where the z-slope increases steeply for moderate values of (*a*) and saturates for higher numbers. Note that the function changes in accordance with the probability density for (*a*) over familiarity (Figure 2), with a strong change over the interval of 2-4 active nodes. The signal to nodes (e.g., *a*) reflects the density of the item representation. A change in (*a*) modulates the representation of the item by making it more or less specific. Thus, the MSDT suggests that item variability is inversely related to density of node activation and performance.

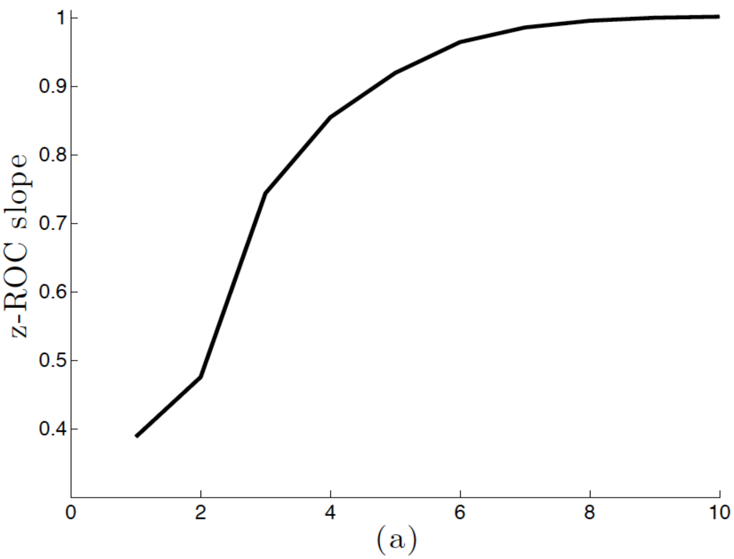
We argue that the empirical value of the z-ROC slope of 0.8 corresponds to a narrow range of plausible number of nodes receiving a signal (*a*). Lower values of the z-slope (e.g., < 0.7) would tend to produce to few active nodes for representing items efficiently (i.e., at least one node needs to be active for efficient representations). For larger values of the z-ROC slope (e.g., > 0.9), the activation of nodes is the result of a wide but shallow allocation of signal. This can be related to sparse and dense representations in neural networks. It is know from previous work that the storage capacity of dense networks is poor (Hertz et al., 1991; Sikström, 2001).

Figure 5. Account of ROC in the MSDT

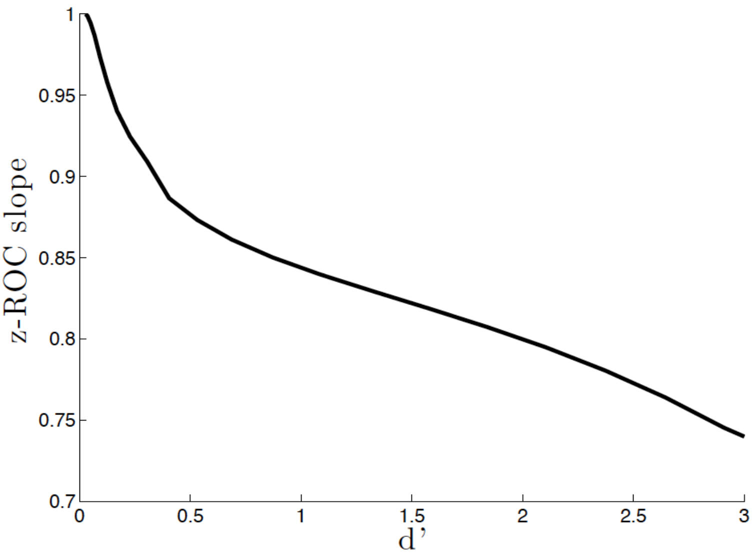
Panel A. z-ROC curves for attentive and inattentive persons



Panel B. z-ROC slope as a function of performance



Panel C. z-ROC slope as a function of (*a*)



Note. Panel A: z-ROC slope for attentive and inattentive where z(H) denotes z-transformed hits, and z(FA) z-transformed false alarms. Red lines represent attentive and blue lines inattentive. Panel B: z-ROC slope as a function of performance for (N) = 80000, when d' = 2 and z-slope = 0.8. Panel C: z-slope as a function of (a).

Decoding of a dense representation is difficult because it involves a large population of neurons, especially in the case of a large ensemble. This additional requirement could affect the performance of encoding a test item. Thus, when several patterns of the representation are superimposed, it gets hard to decode and also introduce ambiguities and interference. Expressed differently, the MSDT suggests that a z-slope neighboring 0.8 is consistent with an efficient item representation, whereas lower or higher values of the z-slope would produce a network with either a highly constrained capability of representing items, or poor storage capacity, respectively.

To summarize, we present a model that extends the framework of signal detection theory, to account for the changes in the old and new item variability of recognition. The model uses a non-linear activation function encompassed by a step-wise function that defines whether a node is active or not. It specifically accounts for the changes in new versus old item variability and the corresponding z-ROC, relates the z-slope to the number of active nodes in the model and provides a unified account of item and response variability. The MSDT also extends the common interpretation of z-ROC data, which holds that the z-slope reflects magnitude of variability in the familiarity distribution. Rather, variation in z-ROC slope is the result of changes in (*a*). The MSDT therefore not only elucidates the need for a formal model of item variability, it also call upon research to investigate which variables affects the input of signal and noise to the underlying nodes in the system to increase the understanding of recognition memory. Further, the model provides a detailed account of memory differences in attentive and inattentive people, with some novel predictions.

First, inattentive persons will exhibit higher z-ROC slopes than attentive. Second, the predicted difference in new and old item variability is to a high degree driven by changes in the new item distribution rather than by an increase in the old item distribution as is commonly assumed in the literature. Third, the model provides a unified account of item variability and response variability, and relates changes in response variability to modulations of the new item distribution.

Current accounts of z-ROC data

A variety of memory models have been proposed to account for ROC data, of which SDT and dual-process theory are highly influential and provides partially different accounts of item variability. The Unequal-Variance SDT (the UVSD) and the Dual Process SDT, the DPSD (Yonelinas, 1994), are shortly described below (1.4.1 and 1.4.2) in terms of the account of the slope of the z-ROC .

The UVSD model

The UVSD assumes that the variance of the old and the new item distributions differ and that the z-slope is equal to the ratio of the standard deviations of the new to the old item distribution, and that old item variability is greater due to a lower proportion of match of items in memory and the presented probe. New items have a standard deviation of 1, and the old item distribution is specified by another parameter, so that performance is determined by *d'*, *c* (the response criterion), and old item variability. The probability of a hit is expressed as (Eq. 6);

$$P(\text{"Old"}|\text{Old}) = 1 - \Phi(\lambda)$$

and the probability equation for a false alarm is expressed as (Eq. 7);

$$P(\text{"Old"}|\text{New}) = 1 - \Phi\left(\frac{\lambda - \mu}{\sigma}\right)$$

where Φ is the cumulative distribution function, λ is the response criterion, μ is the mean of the old item distribution and σ represents the standard deviation of the new item distribution. In other words, the probability of hits is the proportion of the old item distribution above the response criterion, and the probability of false alarms is the proportion of the new item distribution above the response criterion, or (Eq. 8);

$$P(\text{"Old"}|\text{Old}) = (F_o > c),$$

where F_o reflects familiarity for old items. The probability of hits can also be expressed as (Eq. 9);

$$P(\text{"Old"}|\text{New}) = (F_n > c).$$

As the UVSD has one parameter for performance and one for symmetry, the two can be experimentally separated. There is no current consensus regarding the effect of accuracy on z-slope. On the one hand, it has been proposed that the z-slope is affected in a lawful pattern by several manipulations such as word frequency (Arndt & Reder, 2002), list length (Gronlund & Elam, 1994), and divided attention (Yonelinas, 2001), and that these variables also have an effect on accuracy. Accordingly, Glanzer (Glanzer et al., 1999) reviewed the ROC literature and concluded that the z-slope is inversely related to performance. On the other hand, Ratcliff (McKoon & Tindall, 1994) stated that the slope remains constant irrespective of experimental condition and accuracy, this constant value being approximately 0.80. This finding has been replicated several times, showing that the UVSD provides a better account of recognition memory than the EVSD model (Mickes et al., 2007; Ratcliff et al., 1992; Yonelinas & Parks, 2007).

The Encoding Variability Hypothesis

To account for a z-ROC slope below 1.0 it has been suggested that the increase in variance of old items is a result of encoding variability (Wixted, 2007). That is, because the memory strength is selectively increased for some items during study, the old item distribution gets more varied than that of the new items, which results in an asymmetrical ROC with a corresponding z-slope below 1.0. However, the EVH simply states that the difference in old and new item variability is a result of difference in item familiarity, and provides no detailed explanation of how this difference occurs.

There are also other limitations with the encoding variability account. The EVH assumes that the new item distribution is unaffected by encoding variability, which is problematic given the contribution of pre-experimental familiarity on memory. In recognition tests, the most common test materials constitute items that the testee has encountered prior to the experimental test, such as words with varying levels of experiential frequency. Several recent studies have demonstrated that familiarity for test items affect experimental recognition performance, for both words (Estes & Maddox, 2002; Reder et al., 2000), names (Bird, Davies, Ward, & Burgess, 2011; Stenberg, Hellman, Johansson, & Rosen, 2009) and faces (Bird & Burgess, 2008; Bird et al., 2011). Thus, because prior experience with an item modulate the level of evidence for a new item in the test phase of a recognition test, pre-experimental familiarity is related to the new item distribution. It therefore seems essential that models of item familiarity, such as SDT and the EVH, provides a detailed description for why and how variability in the new item distribution occurs, and not only stresses the significance of variability in the old item distribution.

Further, the EVH has been very influential for theorizing of recognition memory decisions, but it has rarely been tested, and when put under scrutiny the outcome has been debated. Koen and Yonelinas (Koen & Yonelinas, 2010) recently tested the EVH by comparing changes in the new to the old item variability (the z-slope) as a result of variations in encoding efficiency. Participants studied two different lists: a mixed list were half of the items were presented for 1 sec., and the other half for 4 sec., and a pure list were all items were studied for 2.5 sec. According to the authors' interpretation of EVH, this manipulation would result in increased new to old item variability in the mixed list. There was no reliable change in z-slope for the two conditions, which Koen and Yonelinas interpreted as inconsistent with the EVH. In two commentary articles (Jang, Mickes, & Wixted, 2012; Starns, Rotello, & Ratcliff, 2012) it was argued that Koen and

Yonelinas conclusions was fallacious because the tested prediction was derived from the mixture-UVSD model, rather than the EVH, and also, that the Koen and Yonelinas study contained both testing related and statistical errors. In a reply (Koen & Yonelinas, 2013), the authors provided an augmented support for their initial finding by some new simulations and armchair reasoning. We have no specific inclination for whether the account of encoding variability or the competing dual process signal detection theory is more accurate, but the described research indicate, at best, a lack of clarity in the SDT account (for further reading, see Rouder, Pratte & Morey, 2010 and Wixted & Mickes, 2010).

The Dual Process Signal Detection Model

In the Dual Process Signal Detection model (Yonelinas, 1994, 2001), recognition memory is based on two different memory processes: familiarity and recollection (*R*). If recognition is based on recollection, the proportion of high confident hits increases. Recognition by familiarity, on the other hand, encompasses a familiarity estimate of the presented item, a process that can be described by signal detection theory. *R* constitutes the proportion recollected items, and items are endorsed as old independently of the response criteria. Items that are not recollected, which occur with a probability of (1-*R*), are recognized with familiarity. The probability of a hit, then, is expressed as (Eq. 10);

$$P(\text{"Old"}|Old) = R + (1\text{-}R) * \left[1 - \Phi\left(\frac{\lambda - \mu}{\sigma}\right)\right]$$

or (Eq. 11);

$$P(\text{"Old"}|Old) = R + (1\text{-}R)(F_o < c)$$

New items are never endorsed with recollection, which is why acceptance of a new item always is based on familiarity according to (Eq. 12);

$$P(\text{"Old"}|New) = F_n > C$$

Because the model includes two memory processes, the account for z-slopes below 1.0 differs from that of the UVSD model. The latter assumes that memory decisions are based on a single evidence variable, and that the higher variability in the old as compared to the new item distribution results in a decrease in z-slope. The DPSD model, on the other hand, relates the change in variance to the two memory processes familiarity and recollection. Because recollection results in a high proportion of correct high confident responses, and familiarity results in both correct and incorrect memory decisions, the variability of the old item distribution is varied. Selective influence of the two processes results in higher (more recollection responses) or lower (less recollection responses) old item variability (Yonelinas & Parks, 2007).

The DPSD model makes a specific prediction on the relation between performance and z-slope. Manipulations that increase recollection more than familiarity (such as semantic versus perceptual levels of processing and full versus divided attention) should result in a decrease in z-slope and increase in performance. But, manipulations that affect both processes equally (such as study duration) should lead to an increase in performance with no affect on the z-slope (Yonelinas, 1999).

The two models have been extensively compared, but no consensus has been reached as to which of the two models offers the best account of recognition memory (DeCarlo, 2002; Onyper, Zhang, & Howard, 2010). Both models can fit asymmetrical curvilinear old-new item ROCs and can account for linear shapes of the z-slope. However, deviations from linearity of z-slopes have been reported for item recognition data. For example, U-shaped z-slopes have been observed for words that had undergone elaborated encoding and words instructed to be remembered from different study lists (Yonelinas et al., 1996) and also, for recognition of complex photographs (Howard, Bessette-Symons, Zhang, & Hoyer, 2006). Further, some studies have reported that z-ROCs are more U-shaped for young as compared to old participants (Glanzer, Hilford, & Kim, 2004; Prull, Dawes, Martin, Rosenberg, & Light, 2006). It has also been proposed that the nonlinear z-slope is a result of random responding, or variability in the response criterion (Malmberg & Xu, 2006; Ratcliff et al., 1994), which indicates the importance of providing a plausible account of such data.

The fact that the DPSD model assumes that recollection is a threshold process, which is necessary for the account of item variability, has been questioned because some studies have indicated that recollection is a continuous process (Glanzer et al., 2004; Qin, Raye, Johnson, & Mitchell, 2001; Slotnick, 2010; Slotnick et al., 2000). On the other hand, the UVSD model must rely on the auxiliary hypothesis of encoding variability to account for the very basic finding that the old item variability exceeds that of the new items, an account that also has been questioned, and which is undetailed and rarely tested. Both the UVSD and the DPSD model assume that the underlying distribution is normal, even though this assumption is difficult to verify empirically (Rouder, Pratte, & Morey, 2010).

We argue that a detailed and elaborate explanation of why the old item distribution is more varied than the new item distribution is needed, and consequently, why the slope of the z-ROC in general is below 1.0 in recognition memory tests.

Attention, dopamine and node activity

According to the MSDT, changes in (*a*) occur as a result of variations in attention, and these variations are related to changes in new to old item variability. Further, behavioral, neurochemical and anatomical studies has revealed that ADHD is related to dopamine. Therefore, the MSDT suggests a relation between item variability and ADHD.

On a neurochemical level, the release of dopamine affects two aspects of the firing probability by tonic (responsiveness to background levels of the environment), and phasic (responsiveness to specific events) components (Floresco & Grace, 2003). Decreases in tonic dopamine affects phasic response efficiency, which results in a strong reactivity to stimulation from the environment, and alterations in tonic dopamine has been related to ADHD (Seamans & Yang, 2004). Further, Ernst and colleagues (Ernst, Zametkin, Matochik, Jons, & Cohen, 1998) compared the presynaptic dopaminergic function in ADHD adults and healthy controls using PET with F18-DOPA, and showed that ADHD participants exhibited lower F18-DOPA values in medial and left prefrontal areas of the brain. They concluded that ADHD symptoms are mediated by dopaminergic dysfunction. This was supported by a recent review of genetic studies of ADHD, where the authors (Bobb, Addington, et al., 2005) concluded that four genes were related to ADHD symptoms, namely the dopamine D4 and D5 receptors, and the dopamine and serotonin transporters.

Further, structures in the brain that has been implicated in ADHD are those with dopaminergic projections from the midbrain (Tripp & Wickens, 2008). Examples of changes in gross brain anatomy are the two findings that overall brain size for ADHD patients is lower than for healthy controls (a reduction that remains into adulthood) (Castellanos & Tannock, 2002), and that a reduction of specific brain areas has been observed in ADHD patients. For example, the caudate nucleus and the globus pallidus, which both contain a high density of dopamine receptors, has been observed to be smaller in patients with ADHD (Swanson et al., 2007). The causative relation between attentional deficits and sub-optimal dopamine levels has also gained support from animal studies. For example, rat studies indicate that burst firing occurs in response to events that are associated with rewards (Mirenowicz & Schultz, 1994) and that dopamine cells respond to appetitive stimuli (Hyland et al., 2002) whereas aversive events inhibit dopamine firing (Ungless, Magill, & Bolam, 2004).

The described differences in excitatory and inhibitory cell activity in attentive and inattentive persons is implemented in the MSDT by manipulation of (*a*). In the former group relative the latter, net input to nodes is higher (reflecting excitatory cell activity) and fewer nodes are activated (reflecting inhibitory cell activity). That is, inattentive persons rely on a broader range of active nodes in recognition, and therefore, on changes in (*t*). However, there is a small activation of nodes because the sum of signal input is equal for attentive and inattentive persons. Because attentional deficits in ADHD have been related to dopamine, it raises the question whether the z-ROC slope may reflect dopaminergic state.

Previous work on computational modeling has showed that individual neurons respond to the environment, which determines the probability of firing after presentation of a stimulus. Such alterations are related to differences in attention. The firing probability, which has been related to levels of catecholamine´s such as dopamine, has been modeled using a sigmoidal activation function (Servan-Schreiber, Printz, & Cohen, 1990; Servan-Schreiber et al., 1998), where a gain-parameter reflects the level of dopamine. When the gain parameter is low, neurons will fire at random and lead to poor cognitive performance and consequently, if the gain parameter is high, cognitive stability will increase and result in high performance.

In the MSDT, the input-output relation in the neural system is described by a stepwise activation function, where a node is active (1) if the input is above the activation threshold and otherwise it is inactive (0). However, the relationship between input and output is modulated by other parameters, in particular the attention parameter (*a*). There is an inverse relation between this parameter (1/*a*) and the input-output relation, and the noise in the system acts by smoothing the activation function. Thus, the input-output relation used in the Servan-Schreiber study can be accommodated within the MSDT by relating high and low attention to differences in a gain-parameter in the non-linear activation function. In Figure 6, node activity is plotted as a function of (*S*). As can be seen, the input-output relation is stronger for attentive than inattentive people, where the old item representation is more sharply activated whereas the new item representation are more distinctly activated. This relationship is modulated by the attention parameter (1/*a*), and can be compared with the gain factor in a sigmoidal activation function.

Interpreting the z-slope as an indicative of dopamine levels

Given the causative relation between attention and dopamine, the MSDT predicts that the z- slope depends on the amount of dopaminergic release. More precisely, the ratio between new and old item variability is expected to decrease for inattentive as compared to attentive persons. To the best knowledge of the authors, no studies have explicitly examined the relation between dopamine levels and the slope of the z-ROC, but some studies show indirect support to this notion. Pertovaara (Pertovaara et al., 2004) observed an inverse relation between pain threshold/response criterion and dopamine binding, and showed that the ROC curve tended to be more symmetrical at lower levels of stimulation, as shown by their Figure 1. The relation between dopamine (D2/D3 receptors) and individual differences in response to painful stimulation was investigated using positron emission tomography to measure the binding potential of dopamine in the brain, and the ROC technique was used to measure sensory discriminability. Further, it has been shown that dopamine D2 receptor binding potential (particularly in the striatum) is involved in central pain modulation, and that the number of D2 receptors is inversely related to tonic levels of pain suppression and lower levels of pain inhibition (Hagelberg et al., 2002).

Further, in a study investigating whether olfactory testing may aid to diagnosis of early PD (Parkinsons disease), because olfactory dysfunction is a typical sign of the disease, ROC curves were compared in PD patients and healthy, age-matched controls. The asymmetry of the ROC curve was inversely related to age (Doty, Bromley, & Stern, 1995). Both aging and Parkinsons disease is related to a negative change in the dopamine system, with a natural decline due to loss of both D1 and D2 receptors in the former (Fearnley & Lees, 1991) and a pathological loss of neurons in the substantia nigra for the latter (Dauer & Przedborski, 2003; Hornykiewicz, 1998).

Experimental test of the models on attention and ROC-curves

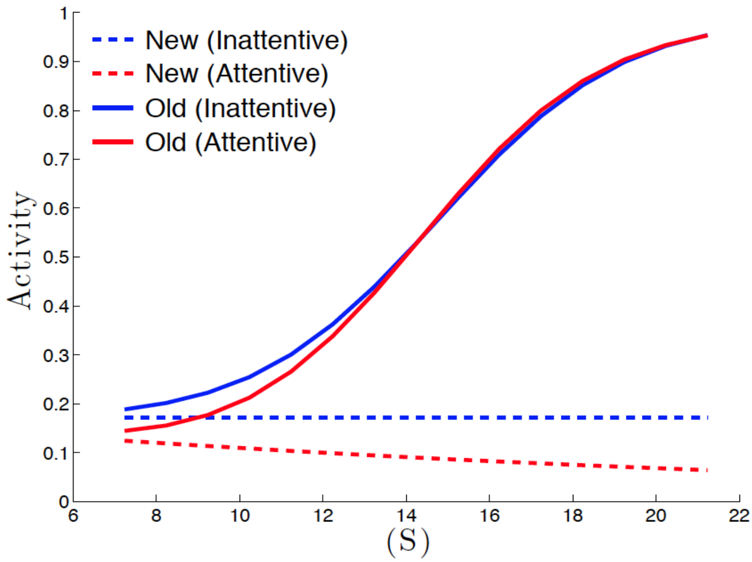
The current experiment aims to test the prediction that the slope of the z-ROC is inversely related to changes in self rated attention, and also, whether the z-slope can function as a potential measure of how behavioral changes are related to variability in recognition decisions. We argue that the MSDT, the UVSD and the DPSD derive different predictions for how the z-slope is affected by differences in attention.

The MSDT predicts that inattentive persons should exhibit higher z-slopes than attentive persons, because according to the model inattentive participants have more diffuse neural representation (i.e., higher values of *a*).

It can be argued that the EVH predicts that the z-slope in inattentive persons should be lower compared to the attentive, because inattentive participants have a higher overall response variability (Castellanos & Tannock, 2002; Gilden & Hancock, 2007) as described above (see 1.2.4). Thus, according to this interpretation, EVH predicts that attentive participants will exhibit a higher z-ROC slope and lower response variability, compared to inattentive people. Further, when the EVH is combined with the idea that encoding is influenced by novelty detection (Metcalfe, 1993), it would predict a z-ROC above one. According to Metcalfe (1993), items with low familiarity are encoded stronger than items with high familiarity, which may lead to smaller variability for old than for new items. Consider for example two items, one that prior to encoding have a low familiarity value (-1) and one that have a high familiarity value (+1). Novelty detectors would encode the low familiarity item more (+2) than the high familiarity item (+1), so that the familiarity strength following encoding is low for the high familiarity item (-1 +2 = +1) than for the low familiarity item (+1 +1 = +2). This would lead to smaller variability for the old items following encoding (+1 versus +2) compared to new items (-1 versus +1). Expressed differently, less familiar items receive increased encoding resources, whereas items with high pre-experimental familiarity receive less encoding. This would lead to diminished variability for the old distribution relatively the new distribution, leading a predicted z-ROC slope that is larger than one.

The prediction derived from the DPSD can be said to be ambiguous. On the one hand, it may be argued that inattentive participants will exhibit spared familiarity and reduced recollection responses, because recollection is more attention demanding. This would result in a higher z-slope for attentive as compared to inattentive participants. However, empirical studies suggest another pattern where familiarity is hampered in inattentive people. For instance, Davidson (et al., 2006) investigated recognition memory performance in Parkinson's disease (PD) patients (a disease associated with sub-optimal dopamine levels) using the remember-know and the process-dissociation procedure to estimate the contribution of familiarity and recollection. The authors observed a decline in familiarity, and not recollection. This is relevant because PD patients suffer from an attentional deficit (Botha & Carr, 2012). Thus, from a theoretical point of view, the z-ROC slope should increase in inattentive compared to attentive people. However, recent empirical data suggests the contrary, revealing that the DPSD does not provide a clear prediction for item variability in attentive and inattentive people.

Figure 6. Node activity as a function of (*S*)



Note. Activity of nodes for old and new items as a function of (S). Solid lines represents old items and dotted lines new items. Attentive in red and inattentive in blue.

The MSDT also provides novel predictions regarding response variability for attentive and inattentive people, namely that the increase in response variability in inattentive is driven by changes in false alarms, whereas response variability in attentive participants is modulated by hit rates.

To test these predictions and compare the three models, we will collect recognition and confidence data on attentive and inattentive participants, and compare the fitting of empirical data to the models.

Method

Participants

A total of 300 participants were recruited from Mechanical Turk (Amazon Mechanical Turk) and conducted the 18 item Adult ADHD Self-Report Scale Symptom Checklist (ASRS). The ASRS is based on the DSM-IV criteria for ADHD, and has been shown to have high validity (Davidson, 2008). Participants with a score > 24 points on part A or part B (criteria for ADHD for adults), or a total score < 17 (high attentive group), were invited to the experiment by a direct link from the questionnaire. Participants with scores between 17 and 24 were excluded from participation in the experiment. Thirty participants (2 women) with a mean age ± SD of 26.9 ± 5.75 years (range 21-36 years) met the criteria for ADHD, and participated in the experiment. An additional forty-five participants (25 women) with a mean age ± SD of 30.6 ± 7.86 years (range 20-42 years), with an ASRS score below the criteria for ADHD, were assigned as controls. The experiment was constructed online with html and JavaScript, and responses were recorded to a MySQL database.

Material

A total of 240 English nouns, describing common inanimate objects with a word length of 3-8 letters, were used as test material. The experiment constituted 3 study-test blocks, with 80 items per block, of which half were used as distractors in the test phase.

Procedure

Written instruction were administered at the outset of the test, requesting the participants to study words in the study phase for a memory test and, at test, to separate studied from unstudied words and to state how certain they are in this decision. Instructions for the encoding and recognition tasks were administered prior to respective phase. In three counterbalanced study-test blocks, each participant was presented with 40 words during study and 80 words at test, half of which were new. At study, each word was presented during 2 sec., centered on the screen in black font on white background, and the task was to remember the word for a subsequent memory test. The test phase followed immediately after the study phase, where participants were presented with previously studied items intermixed with new ones. The task was to differentiate old from new words, and if accepted as old, give a confidence judgment for the response on a 6-point scale. A confidence response of 1-3 equaled a low to high degree of certainty that the item was old, whereas a response of 4-6 corresponded to low to high confidence that the word was new.

Each word was presented slightly above the center of the screen, in black font on white background. Responses were given by clicking response circles presented below the test word, matching old/new responses and confidence interval, using the mouse. Both responses were self-paced. Conducting the ASRS questionnaire took approximately 5 minutes, for which each participant was rewarded 0.25 USD. The memory test took approximately 25 minutes, which was rewarded with 3 USD.

Results

Averages for the experiment can be seen in Table 1. Hit rates did not differ significantly between the two groups [$F(1,74)=3.49$, *ns*.], although there was a tendency ($p<0.067$). There was a reliable difference in both false alarms [$F(1,74)=21.34$, $p<0.001$] and d' [$F(1,74)=20.74$, $p<0.001$]. There was a difference in response bias ($F(1,74)=14.52$, $p<0.001$), where inattentive participants made liberal memory judgments and attentive exhibited a conservative response bias.

Slope of the z-ROC

To test whether z-slope differed between the two groups, two outliers (one in the attentive and one in the inattentive group) were excluded due to extreme z-slope values (> 4.0). This resulted in a total of 73 out of 75 participants (29 in the ADHD group and 44 in the healthy group). There was a reliable difference as to z-slope [$F(1,72)=8.97$, $p<0.005$]. We also calculated d_a, which is known to be a more robust estimate of sensitivity when the variability of the new and old item distributions vary (Macmillan & Creelman, 2005). The analysis revealed a significant difference between the two groups [$F(1,72)22.51$, $p<0.001$], confirming the analogous analysis with the d' measure.

Due to the disproportionate low numbers of female inattentive participants, we balanced the groups as to gender by removing the female participants from both groups (resulting in 28 participants in the inattentive group and 20 in the attentive group) to control whether the difference in z-slope was a result of gender inhesion in the analysis. As the observed differences between groups regarding both performance (F=20.9, p<0.001) and z-slope (F=14.68, p<0.001) remained, we concluded that the difference in z-slope was not affected by gender.

To control that the difference in z-slopes was not a result of an accuracy confound, we divided the data according to performance and attentiveness and investigated the effect of attention and performance on z-slopes. This resulted in one group with 22 participants (d'<1.5) and one with 37 participants (d'>1.5). There was an effect of group [F(1,58)=5.19, p<0.05], i.e., a difference in z-slope for attentive and inattentive, but not of performance [F(1,58)=2.76, *ns*.]. To corroborate this analysis, we compared the whole groups as to z-slope, entering performance (d') as a covariate and group as a fixed factor. This revealed a reliable effect of group (F(1,72)=6.34, p<0.05), but not of performance F(1,72)=0.004, *ns*.). We interpret this as that the difference in z-slope is not a result of changes in performance.

Table 3. Performance, bias and z-ROC slopes for attentive and inattentive

Measure	Groups			
	Inattentive		Attentive	
	Mean	St. D	Mean	St. D
Hits	0.75	0.17	0.81	0.11
False Alarms	0.38	0.25	0.16	0.11
C	-0.19	0.48	0.14	0.27
d'	1.12	0.93	2.16	0.90
da	1.19	1.18	2.54	1.19
z-slope	0.82	0.33	0.57	0.36

Note. d', A' and z-ROC slopes are calculated in accordance with Macmillan & Creelman (2005). St. D denotes standard deviation.

Response variability

Alterations in response variability in attentive and inattentive people (Gilden & Hancock, 2007) are interesting because it can reveal underlying aspects of differences in performance, and because the MSDT predicts that inattentive exhibit higher response variability than attentive people. To investigate these differences, we compared the standard deviation for old and new responses (hits and false alarms) over subjects, as well as for the response criterion and performance for the two groups, for which averages are reported in Table 3. As can be seen, the inattentive participants have a larger variability as compared to attentive participants, and the difference is higher for new than old items as predicted by the MSDT. When tested with the F-test for equality of two variances (Snedecor and Cochran, 1983), there was a reliable difference for hits [F=2,39, p<0.01], false alarms [F=4,73, p<0.001], the response criterion [F=3.16, p<0.005], but not performance (d'), over the groups.

Because the MSDT predict higher response variability in inattentive participants due to changes in new item responses, and lower response variability for attentive due to changes in hit rates, we compared standard deviations for old and new items within the groups. Indeed, standard deviations for hits were significantly lower than standard deviations for false alarms for inattentive participants [F=2,16, p<0.05], but not for attentive [F=1,0, *ns*.].

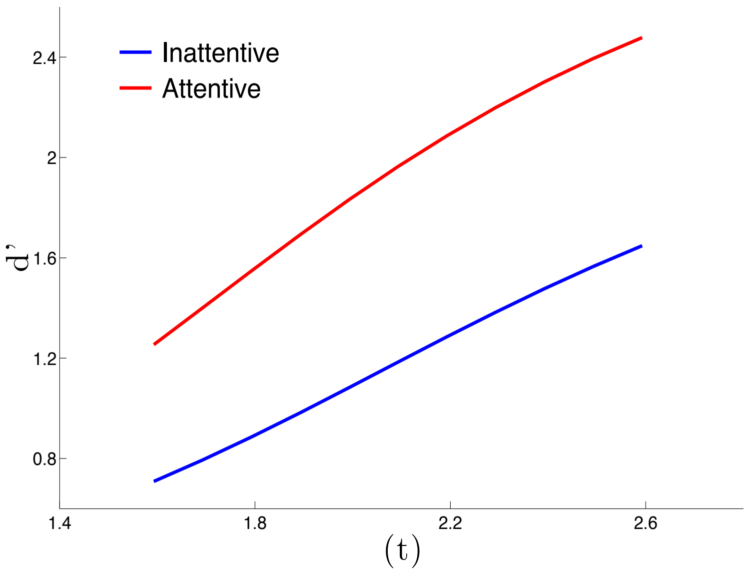
The MSDT also predicts that attentive and inattentive participants implement different response bias, where the latter group should exhibit a more liberal response criterions relative the former. Figure 7 plots performance as a function of (t). Inattentive people exhibit a lower performance rate over changes in (t), and a planer increase in performance. This must be considered in relation to how old and new item variability is influenced by changes in (t) (see Figure 5, panel B). For inattentive persons, old item variability increases moderately whereas the new item variability increases steeply and thereafter decreases steeply within reasonable values of (t). For attentive persons, old items have a low variability over a high interval of (t). Thus, performance is more sensitive to changes in (t) for inattentive than attentive participants, leading to a higher variability for inattentive than attentive participants. Indeed, there was a reliable difference in (C) between the two groups (F(1,74)=14.52, p<0.001), where inattentive implemented a more liberal response bias than attentive.

Model fitting

In a quantitative comparison of the three models, we fitted each model to hits, false alarms and z-slope values with MLE, and the mean-square error (*mse*) is reported, which is a risk function that estimates the squared error loss of an estimator. Given than the density function is known the MLE approach provides reliable parameter estimation in model comparison as compared to least-square estimation, and determines the parameter value that corresponds to the probability distribution that makes observed data most likely. Hits and false alarms were fitted with a binomial distribution corresponding to p(old) and p(new), using MLE. However, for the z-ROC slope, a normal latent distribution was used to compute MLE, using the z-ROC value and the standard deviation for the z-ROC. When fitting the UVSD, normal unequal-variance distributions are implemented for both probabilities and z-ROC slopes. For the DPSD model, hits, false alarms and z-slopes are generated with a normal distribution with equal variance and a threshold.

The models were fitted to group data and to individual data, and two different fits were performed: a 4-parameter solution and a 3-parameter solution. The parameters were fitted to 120 data points (old and new responses).

Figure 7. Performance as a function of (*t*)



Note. Performance as a function of (t). Attentive in red and inattentive in blue

Fit to group data

In the 4-parameter solution, the following fitted parameter settings were generated. For the MSDT, (*a*) was set selectively for attentive (*a*=2.25) and inattentive (*a*=3.42), a common value for (*t*) occurred for both attentive and inattentive (*t*=10.41). The performance parameter was the same for both groups (*S*=10.68). For the UVSD, performance was higher for attentive (d'=1.13) than inattentive (d'=0.57), and the old item standard deviation (denoted σ_O in the context of analysis) equaled 0.61. The threshold was set to 0.82, and was used for both attentive and inattentive. For the DPSD model, both familiarity and recollection contributes to recognition, where familiarity is represented by d', whereas the threshold process recollection is represented by *R* (i.e., the adjusted standard deviation induced by recollection). Performance for attentive and inattentive was d'=0.85 and d'=0.57, respectively, and the difference in old item variability (induced by *R*), was set to 0.14. The threshold was 1.21.

Table 4. Predicted, observed and parameter estimation values for performance and z-ROC slope for the three models in the group fit

		Model						
			MSDT		UVSD		DPSD	
Meas.	**Group**	**Data**	**3**	**4**	**3**	**4**	**3**	**4**
H	*Attentive*	0.81	0.83	0.76	0.83	0.82	0.87	0.88
	Inattentive	0.75	0.67	0.80	0.75	0.74	0.74	0.71
FA	*Attentive*	0.16	0.16	0.14	0.29	0.27	0.31	0.26
	Inattentive	0.38	0.33	0.38	0.29	0.27	0.31	0.34
z-slope	*Attentive*	0.57	0.57	0.66	0.58	0.57	0.62	0.66
	Inattentive	0.82	0.80	0.74	0.82	0.81	0.82	0.81
MLE			113.1	113.1	115.6	115.4	117.1	113.1
mse			0.001	0.001	0.004	0.004	0.005	0.004
BIC			0.56	4.01	0.85	4.20	0.60	4.17

Note. Meas. denotes measure. Att. and Inatt. stands for Attentive and Inattentive. H and FA denote Hits and False Alarms. 3 and 4 represents the 3-parameter and 4-parameter solution, respectively (see text for details). MLE denotes Maximum-likelihood estimation and mse denotes the mean-square error of the fit. Data represents empirical data from the experiment.

The parameter values were optimized using MLE to fit the data (hit rates, false alarm rates with a binomial distribution and z-slopes using a normal distribution) for both implementations, so that values for the variability parameters for UVSD and DPSD were accommodated for empirical data. Table 4 presents the result for all group fits. The

MSDT provides a marginally better fit than the UVSD and the DPSD models, as the MSDT generates a lower error and MLE value. We also computed the Bayesian information criterion (BIC) for each model and implementation, computed as $BIC = -2 * \ln(L) + k ln(n)$,where L is the maximized value of the likelihood function for the model, k is the number of parameters and n is the number of observations. BIC values are reported in Table 4, which indicated a better fit for the MSDT.

Table 5. Averages for MLE, mse and BIC for the fit to subjects

		Model					
		MSDT		UVSD		DPSD	
Parameter	**Group**	**3**	**4**	**3**	**4**	**3**	**4**
(S) /d'	*Att.*	2.61	21.94	1.04	1.17	1.05	1.01
	Inatt.			0.59	0.57	0.58	0.5
(σ_o) / (a) /R	*Att.*	11.89	3.96	0.80	0.67	0.25	0.17
	Inatt.	20.24	12.8				
(t)	*Att.*	0	8.41	0	0.83	0	1.22
	Inatt.						
MLE		144.2	143.3	124.4	177.7	141.1	132.2
mse		0.03	0.03	0.02	0.01	0.03	0.02
BIC		4.30	8.47	4.13	8.74	4.08	8.11

Note. Att./Inatt. denotes Attentive and Inattentive. 3 and 4 denotes the 3 and 4 parameter solutions for each model. (S) and d' denotes the signal in the MSDT and performance in the UVSD and the DPSD models. *(σ_o) / (*a*)* /R denotes the standard deviation of the old item distribution (the UVSD), the attention parameter (the MSDT) and the Recollection component in the DPSD model. (t) denotes the activation threshold (the MSDT), or the response criterion in SDT. MLE is the Maximum-likelihood estimation value for each model and implementation, *mse* denotes mean-square error of the fit, and BIC is the Bayesian Information Criterion. All values are computed as averages for n=31.

Looking at predicted and observed values for the data, the MSDT predicts false alarms for both groups more accurately than both the UVSD and the DPSD model, whereas hit rates fort both groups and z-slope value for inattentive were more accurately predicted by the UVSD. As for inattentive z-slopes, the MSDT and the UVSD predicted approximately similar values, with a marginal difference in favor of the UVSD model. Predicted z-slope values were more accurate because the UVSD has a separate variance parameter for the old item distribution, whereas the MSDT uses a parameter, (*a*), reflecting node activity that modulates variability of the new and old item distribution.

In the 3-parameter fit, we used a 3-parameter solution where the threshold was set to 0 (i.e., an unbiased threshold). In the MSDT, (*a*) for inattentive and attentive was 3.35 and 2.20, respectively. Performance (*S*) equaled 10.47. For the UVSD model, performance was 1.11 for attentive and 0.58 for inattentive, and σ_o reached 0.82. In the DPSD model, d' equaled 1.01 for attentive and 0.56 for inattentive. *R* equaled 0.14. The results of the fit can be seen in Table 4. The MSDT generated lower MLE and mse values, and predicted false alarm more accurately for both groups, and z-slopes for attentive participants. The UVSD made better estimates of hits for attentive and inattentive, and a more accurate z-slope value for inattentive participants. The BIC values indicate that the MSDT provided a somewhat better fit.

Fit to subject data

The model was fitted to individual responses and z-slopes using 120 data points using MLE and *mse*. Averages for MLE, *mse* and BIC, as well as fitted parameter values over the participants are reported in Table 5. As can be seen, MLE are lower for the UVSD than the DPSD model and the MSDT. As for BIC values, the DPSD seems to provide a marginally better fit. We compared each model on these values, as well as the implementations. There were no significant differences as to MLE, *mse* or BIC between the MSDT, the UVSD or the DPSD model for neither the 3-parameter nor the 4-parameter solution (all *p*s > 0.1). When BIC values were compared between the 3-parameter and the 4-parameter solutions, the MSDT, the UVSD and the DPSD models exhibited reliable differences [F=50.3, 57.2 and 40.6, respectively, all ps < 0.001].

Discussion

The present experiment was set up to investigate the prediction of the MSDT, i.e., whether differences in attention (i.e., comparing participants with high or low ASRS score, where high scores is indicative of ADHD) would generate differences in the new to old item variability. The prediction is based on the link between (*a*) and the variability of the old and new item distributions. Indeed, there was a difference in z-slope, as predicted, showing that inattentive persons exhibit a lower ratio of new to old item variability as compared to attentive persons. According to the MSDT, this group differences in z-slope occur because inattentive exhibits a higher new item variability, which diminishes the difference in new to old item variability.

The MSDT also predict that increased response variability in inattentive people is the result of changes in responses to new items, whereas the analogues effect in attentive people occurs because there is a change in endorsement of old items. Indeed, there was a reliable difference in both response criterion and responses to old and new items between the groups.

Attentive and inattentive differed considerably in performance, and also, a considerable difference in response bias. Whereas hits were roughly approximate, the inattentive group exhibited a greater proportion of false alarms than the participants in the attentive group. This could be interpreted as that the difference in performance also affected the z-slopes, as the new to old item variability is related to response bias. However, when z-slopes were compared across constant levels of performance (d'), the difference in z-slope remained. The difference in performance was predicted by the MSDT, and is related to changes in (*a*). Thus, inattentive are more responsive to changes in (*t*) variability than attentive because the inattentive group exhibits an allocation of nodes where a low number of active nodes receives an input of signal.

General Discussion

In the present paper, we have introduced the MSDT, a model that offers a multidimensional extension of SDT, where familiarity is a sum of non-linear activations of nodes and where the underlying distributions are binomial. Cognition, and especially perception, has been described with multidimensional models albeit with a focus different from that of the MSDT. For instance, a previous modulation of perception has succeeded in differentiating the perceptual system into subcomponents, by relating SDT parameters to perceptual independencies using General Recognition Theory (GRT). For instance, perceptual independence of a blue square means that perceptual processing of one of the two components (the color blue) is unrelated to processing of the other (the square). Thus, GRT can reveal interactions of percepts created by physical stimulus features, and be estimated by relating perceptual space to item responses (Ashby, 1992). MSDT differs from GRT because it is fundamentally based on the idea of a larger number of features or dimensions. The prediction of z-ROC slopes in the MSDT requires a larger number of dimensions, and is not directly applicable to representations including, for instance, two dimensions.

Further, the MSDT provides a new perspective on how to account for item variability in recognition memory, encompassing changes in responses to studied and unstudied items, variations in response bias, the relation between performance and the new to old item variability. Importantly, it provides a new perspective of how to interpret and use the slope of the z-ROC were attention affects performance, response variability and item variability. We have elucidated the limitations with the current most popular accounts of item variability (the UVSD/EVA and the DPSD models) and thereby highlighted the need for an extensive, formal model of item variability.

Model comparison

The quantitative comparison of the three models revealed that when the models were fitted on group level, there was a small difference in MLE, *mse* and BIC in favor of the MSDT. However, when the models were fitted to individual data and the difference in MLE, *mse* and BIC averages for the models were tested, it was showed that neither of the models provided a reliably better fit to data. However, the models differ more substantially in other respects.

The MSDT differs from the UVSD and the DPSD concerning the account of performance and item variability. In SDT, two latent familiarity distributions with different variability reflect old and new items where the placement of a response criterion (C) distinguishes correct from incorrect responses, and the difference in the mean of respective distribution defines performance (d'). Because the new item distribution is assumed to have a variability of 1, and the old item distribution variability increases with encoding, z-ROCs are expected to exhibit a slope below 1.0. Indeed, the majority of empirical z-ROC curves display such z-slope values (Glanzer et al., 1999; Yonelinas & Parks, 2007). Thus, performance is the result of the placement of the response criterion and the z-ROC slope varies depending on the variability of the old item distribution. However, the UVSD must rely on an auxiliary hypothesis, the EVH (Wixted, 2007), to explain why the old item variability is greater than that of new items, and only relate item variability to differences in item familiarity. The EVH does not specify which variables and how these variables affect encoding variability or retrieval variability, nor does it provide a formal account of item variability (for a description of the limitations with the item variability account in SDT models, see Ratcliff & Starns, 2009).

The DPSD provides a different account of performance and item variability because the model is a hybrid of SDT and high-threshold theory, and assumes that a recognition decision is based on two different retrieval processes: familiarity and recollection. The former is described as a signal detection process similar to that of the EVSD, and assumes that evidence is equally variable across targets and lures. Recollection results in correct responses and have no influence on inaccurate responding. Thus, recollection is described as a threshold process and the two processes in combination can produce z-ROC slopes below unity because the SDT process results in equal variance whereas recollection and familiarity in combination decreases the variability of the old item distribution. However, the DPSD accounts for the differences in item variability given that recollection is a threshold process. We do not wish to elaborate on whether this description is correct or not, but we regard the account as perilous because the threshold assumption is questioned (Glanzer et al., 2004; Qin et al., 2001; Slotnick, 2010; Yonelinas, 1994, 1999).

Another important feature of accounts aimed at explaining ROC data is the understanding of how changes in performance affect the z-slope. This has been quite extensively investigated (Arndt & Reder, 2002; Ratcliff et al., 1994; Rotello et al., 2004). Whereas the UVSD does not provide a detailed explanation of this relation, the DPSD assumes that variables affecting familiarity and recollection differentially also has different effects on the z-slope, whereas manipulations with analogous effect on both retrieval processes has no effect on the slope of the z-ROC (Yonelinas & Parks, 2007).

According to the MSDT, variability in the activation threshold and the number of nodes receiving signal affects performance, which is attenuated in inattentive as compared to attentive persons. The model accounts for both equal and unequal variability in the familiarity distributions because changes in the distribution of signal and noise affect the variability of both the new and old item distribution. Thus, the model is fully described by three parameters: signal (*S*), the activation threshold (*t*), and the input of signal to nodes (*a*). The interaction of these parameters has been described in the introduction (see 1.2.1), and will be elaborated below.

The MSDT is similar to conventional SDT in several respects. First, memory performance is conceptualized as the result of the comparison of a probe and memory traces in memory, where each presented test item induces a certain level of activity in the nodes that corresponds to the features of the item. However, for every recognition decision, (*N*) nodes are provided with input of both signal and noise or noise only, and because only a sub set of these nodes gains an input exceeding an activity threshold (*t*), the number of active nodes that contributes to recognition varies according to (*a*) and (*t*). Thus, the MSDT is a multidimensional extension of SDT because item variability and performance is determined by allocation of signal and noise to *N* nodes (*a*) and a threshold (*t*), which acts in a way similar to old item variability and the response criterion in SDT. Encoding manipulations such as that of strength variables (study time, elaborate encoding etc.) are expected to affect performance as predicted by SDT.

Even though the parameters of the MSDT are similar to those of the competing models, they differ in terms of description level. In both the UVSD and the DPSD models, the parameters manipulate the latent familiarity distributions by modulating the distance between and the variability of the distributions, and the placement of a response criterion. However, in the MSDT the analogous changes occur by manipulating node activity. This is an important difference of the two models, being both an advantage and a possible limitation. On the one hand, the MSDT natively describes several variables related to, in this case, attention, but also specificity of the neural basis of the memory representation. The MSDT therefore provide an extended account of recognition memory, where performance and bias are delineated in terms of both strength variables and attention. On the other hand, because changes in (*a*) leads to changes in both performance and response bias, it is hard to establish how variables affecting (*a*) differentially affects performance and bias. This caveat may be considered in light of the advantage of using a parameter (*a*) estimating both neural activity as well as the psychological processes underlying performance and bias, thereby providing a causative chain over neural activity, psychological processes and behavior.

Another interesting aspect of the MSDT regards the conceptualization of features (Malmberg, Steyvers, Stephens, & Shiffrin, 2002). Here, each node represents a feature of the stimulus. However, a feature can represent different levels of abstractness of the stimuli, such as physical character of a word, or a semantic aspect of the stimuli. Thus, in the MSDT, a feature is a subset of an item that bears meaning.

Because the probability of active nodes is described by a binomial and not a normal distribution, the depiction of probability yields additional information with implications for the understanding of recognition memory data. The MSDT suggest that the relation between the ratio of the new to the old standard deviation and the z-slope is inaccurate, and that the linearity of the z-shape cannot be used to distinguish a Gaussian distribution from a binomial distribution. The latter may be important for the understanding of the shape of the z-slope. The non-linear z-shape has been interpreted as evidence for an additional retrieval process (Yonelinas et al., 1996), and as the result of random responding and variability in the response criterion (Malmberg & Xu, 2006). The use of a binomial distribution is also relevant because it relates to the recent interest in questioning whether normal distributions are good approximations to underlying distributions, which has lead to a renewed interest in Bayesian inference. In Bayesian analysis, the probability of success is estimated by taking into account the prior belief of a specific outcome of the event (Kruschke, 2010). Using the standard Bayesian method, where the prior distribution is included, additional information is encompassed in the analysis, which reveals different information about the data as compared to traditional null hypothesis significance testing. A similar approach is taken in the MSDT. By describing the distribution of active nodes as dichotomous, the MSDT asserts the implications for probability of success in a recognition test that differs from that when a normal distribution is used. Thus, by describing data with a binomial distribution, the MSDT provides a different perspective on recognition memory data as compared to models using Gaussian distributions. We suggest that future studies on item recognition should investigate the z-shape based on Gaussian and binomial distributions for those manipulations that previously have induced non-linear z-slopes.

Relating changes in node activity to changes in dopamine

In the present study, we were interested in the prediction that changes in z-slope can be related to differences in attention. We suggest that the slope of the z-ROC reflects how item variability is related to behavioral and biochemical changes, more precisely, to dopamine levels. This relies on an assumed relation between changes in (*a*) and dopaminergic state, and that dopaminergic changes mediates ADHD symptoms. The functional relation between dopamine and ADHD symptoms has been well investigated and supported (Ernst et al., 1998; Seamans & Yang, 2004; Solanto, 2002; Volkow, Fowler, Wang, Ding, & Gatley, 2002). Support for the dopamine account of attentional deficits has been reported from studies of functional neuropsychological (Ernst et al., 1998), biochemical (Bobb, Addington, et al., 2005; Bobb, Castellanos, Addington, & Rapoport, 2005; Seamans, Floresco, & Phillips, 1998), and animal studies (Mirenowicz & Schultz, 1994; Ungless et al., 2004).

The MSDT suggests that the z-slope can be conceptualized as an indirect measure of how tuned neurons are to a certain stimulus. Sparseness of neural representations is inversely related to (*a*) and the z-ROC slope. The notion that verbal (and visual) material is characterized by sparse coding finds support in the literature. Recently, Quiroga and colleagues (Quiroga et al., 2008; Quiroga et al., 2005; Viskontas, Quiroga, & Fried, 2009) studied encoding density for pictures and written names of famous persons (i.e., Halle Berry, Jennifer Aniston), and showed that certain neurons exhibited selective firing for the face and the written name (Quiroga et al., 2008). In the MSDT, high values of (*a*) (e. g., z-slope > 0.9) imply a dense neural representation of the stimulus. Dense encoding is relatively uncommon, but has been observed in the rat olfactory system (Vincis, Gschwend, Bhaukaurally, Beroud, & Carleton, 2012), and unpublished data has supported this view on human olfactory processing by accounting for high z-slopes with the MSDT (Hellman et al., submitted).

Servan-Schreiber (et al., 1990; et al., 1998) developed a standard connectionist network model where the release of dopamine was directly related to the slope of a logistic activation function (e.g., a sigmoidal curve). According to the model, the increased responsivity mediated by release of dopamine is simulated by an increase in the gain parameter, and the activation function - determined by the gain - in turn modulates the activation of units as a function of net input. This is visualized in their Figure 2 (Servan-Schreiber et al., 1998), where it can be seen that for increasing levels of dopamine, the slope of the logistic function increases. The Servan-Schreiber study provided a plausible account for the relation between dopaminergic functioning and the gain parameter. The findings reported in the present paper can add support to these findings by showing that attentive and inattentive persons differ in the variability of the underlying familiarity distributions as a result of different application of the non-linear activation function (which, in turn, is determined by the gain parameter). We argue that this difference occurs because participants with high attention maintain a non-linear activation function, which results in an increase in old item variability, whereas participants with attentional deficits shift between a linear and non-linear activation function (i.e., variability in the gain-parameter), and generates lower net old item variability.

It should be noted that the suggested relation between dopamine and z-ROC slope is, as discussed above, indicating. A causative relation between a measure of item variability and dopamine level assumes that the above described connection between attention and ADHD, ADHD and dopamine and the z-ROC slope and attention is unequivocal. That cannot be verified given the empirical data reported here. However, the suggested relation between item variability and attention is reasonable and novel, and implies that attentional deficit is related to both response variability and item variability quite different from that of a healthy population.

Other models of item variability

A potential problem with the DPSD model constitutes the conceptualization of recollection as a threshold process, where the recollective threshold is thought of as differing high from low levels of episodic retrieval. However, some other dual process models relevant for the research topic have been developed, which does not necessitate a threshold process, namely; the two-dimensional SDT model (Glanzer et al., 2004), the sum-difference theory of remembering and knowing (Rotello et al., 2004), the some-or-none model (Kelley & Wixted, 2001) or the mixture model (DeCarlo, 2002). Even though we do not whish to accomplish an elaborate comparison of the MSDT and these models, or enter the fervent debate of single- and dual process theory, it holds a central place in the recognition memory literature. We will therefore described these models shortly below.

In the two-dimensional SDT (2DSD) model (Glanzer et al., 2004), item and recollection based judgments are based on the same latent distributions, resulting in an increase in item memory when source memory performance is elevated, and consequently, higher source memory performance for an increase in item memory performance. One memory strength dimension has been devoted to both sources, meaning that strength and variance can vary for the two sources. To distinguish item memory performance, a response criterion is placed between the new item distribution and the source distributions, whereas applying a response criterion between the two source distributions delineates source memory discrimination. The 2DSD assumes Gaussian latent distributions, resulting in ROC predictions similar to those of the UVSD (i.e., both source and item ROCs are curved in p-space and linear in z-space). The free parameter constraining the variance of the old item distribution produces z-slopes below 1.0, and enables a dissociation of asymmetry and performance. The 2DSD model have contributed to the understanding of source memory recognition, but provides no novel predictions for ROC data. The model does not offer a clear definition of how item and source recognition is related, because this depends on the relation between the two source dimensions and the type of source information used in the particular decision.

The sum-difference theory of remembering and knowing (STREAK), which is a two-dimensional signal detection model that accounts for item and remember/know data, has one dimension devoted to global familiarity and another (orthogonal) dimension that represents recollection (Rotello et al., 2004). The latent distributions for familiarity and recollection have equal variability, but the old item variability is greater for old than new items. What differs STREAK from other dual process models is the separated delineation of remember/know responses and item responses. Remember and know responses in combination are parallel with the confidence ROC as these judgments encompass all items above the old/new response criterion, whereas remember ROC points can vary along the confidence ROC. Thus, the remember/know ROC slope can be different than the standard item memory z-slope. However, the novel aspect of STREAK is also the models major weakness. STREAK provides new predictions for remember/know data, but the use of the remember/know paradigm as a way to estimate the contribution of familiarity and recollection has received considerable and convincing critique (Dunn, 2004, 2008; Malmberg & Xu, 2006).

The mixture model, proposed by DeCarlo (DeCarlo, 2002, 2003), has been used to account for item and source recognition ROCs. Item memory is described by Gaussian memory strength distributions with equal-variance, and an attention process results in increased memory strength for some (attended) items. New items encompass a normal distribution, whereas old items generate a mixture of two equal variance distributions, where one distribution is shifted to the right because some items are more attended (i.e., has an increase in memory strength). Thus, the attention process influences item recognition differentially, in accordance with the encoding variability hypothesis (Wixted, 2007), but avoids the theoretical problem of negative memory, as described by DeCarlo (2002). That is, if some items are associated with an increment in memory strength during study, others consequently decrease in familiarity. Because the underlying distributions are assumed to have equal variability, the ROC is predicted to be asymmetrical in probability space, and due to the selective increase in memory strength for attended items, the model allows linear z-ROC slope below 1.0. The mixture model may seem similar to the MSDT, because an attention process modulates the new to old item variability. However, in the mixture model, attention is not clearly defined, and if the attention parameter modulates the new to old item variability, the effects thereof should be described in detail. But to determine the effect of attention manipulations at encoding on the z-ROC is difficult because attention operates rather non-monotonically. In the MSDT, attention is viewed differently, namely as an endogenous variable that divides the data set. That is, the MSDT accounts for changes in attention as a function of changes in (*a*), which in turn is related to chatecholaminergic state. Further, the MSDT provides several new predictions of both item and response variability, which is not the case in the mixture model.

Another theory that may seem similar to the MSDT is the attention-likelihood theory (or the ALT), presented by Glanzer and colleagues (Glanzer & Adams, 1985, 1990; Glanzer, Adams, Iverson, & Kim, 1993). ALT shares with the MSDT the assumption that the latent distributions are binomial, and that attention modulates recognition performance. In essence, the theory assumes that stimuli consist of *N* number of features that are either marked or unmarked, where a marked feature represents that it was present in the study list. At study, features are sampled and marked, and the proportion of features sampled for the presented stimulus depends on attention. However, also new items can have marked features due to random noise in the decision process. If the noise level exceeds zero a proportion of the new items will be marked with a certain probability, and the probability of a marked feature for an old item simply equals the number of marked features divided by the total number of features (*N*). At recognition, the features are sampled, and the number of marked features is binomially distributed. The recognition decision is based on the estimated difference of the sample size and the number of marked features for each stimulus. Thus, the old-new response is based on a log-likelihood ratio given the class of stimulus (i.e., the amount of attention devoted to the stimulus).

ALT has been criticized for being complex. To make a recognition decision for a presented test item, the log-likelihood ratio must be estimated. If the comparison stands between an old and new item (for item A), two log-likelihood ratios are needed. However, if the experiment uses two classes of stimuli, the log-likelihood comparison will encompass six different possibilities, namely a comparison of sampled features for old and new items for: A(old)/A(new); B(old)/B(new); A(old)/B(new); B(old)/A(new); A(old)/B(old) and B(new)/A(new). To estimate the log-likelihood for each stimulus, a second set of binomial distributions must be constructed. That is, not a binomial for *x* (the dependent variable) but for the comparison of new and old for the different classes of stimuli (for a more detailed discussion on the topic, see Murdock, 1998, and for a similar critique, see Hintzman, Caulton and Curran, 1994). In addition, the ROC account of ALT has been questioned. In the context of the mirror effect, ALT holds that participants use the same likelihood ratio for items in both the strong and weak condition, which leads to a situation where the response criterion and the criterion for confidence shifts in different directions over different levels of familiarity. In essence, this put high demands on the participants, because she must be aware of the location of both the target and lure distributions as well as their mathematical forms. For a detailed description of the ROC account of ALT, see Wixted and Gaitan (2002).

Prior knowledge and recognition

The MSDT stresses changes in the new item distribution for both response variability and item variability, which implies that the model encompass the contribution of prior knowledge on experimental recognition tests. Neither the UVSD nor the DPSD acknowledges that the new item distribution is important for the understanding of item variability, but rather focuses on changes in the old item distribution. As proposed by the

MSDT, the new to old item variance distribution is affected by changes in new item responses. We argue that this difference occurs because many test items (i.e., words) are familiar at encoding, which affects the reinstantiation of the probe at recognition differently for previously known and unknown items. Therefore, the model should account for previously described phenomena related to pre-experimental familiarity, namely word frequency effects. According to the MSDT, high frequent items are associated with more pre-experimental contexts than low frequent items, which results in an increased variability in the input to nodes containing information about the context associated with the item. The increased variability, in turn, lowers familiarity for the word, which makes it difficult for the testee to distinguish pre-experimental from experimental contexts related to the word. Whereas high levels of context related variability decreases (i.e., induced by an increment in the number of contexts), the expected value of the input is equal for both high and low frequent items (Sikström, 2001).

The MSDT also imply a fusion of two commonly separated fields of research; item variability in recognition memory and ADHD symptomology. We propose that these two fields can contribute to each other because a measure of item variability commonly used in the recognition memory literature may in fact reflect differences in attentional skill, and possibly differences in dopaminergic state. Thus, future studies should relate both attentional differences and other behavior related to dopaminergic state to item variability to increase the understanding of both memory processes and psychological aspects of dopaminergic state.

In summary, we present a model that on the one hand accounts for a wide interval of item variability distributions as well as the relation between the z-slope and performance. The model also encompasses the common finding that response variability is higher for persons with attentional deficits and explains why inattentive people perform worse than attentive on recognition memory tests and why response variability differs in attentive and inattentive people. The MSDT accounts for several characteristics of recognition memory, namely memory strength, item variability, response variability and ROC data. It extends the understanding of item variability by relating the z-slope to different manipulations, neural density, catecholaminergic state and differences in attention. It provides a formal theoretical basis for these phenomena, and sets up several novel predictions for future research. Future research on z-ROC curves should consider binomial rather than normal latent distributions underlying recognition decisions, relate the effect of changes in new and old item variability on z-slope and performance and consider whether the variables affecting z-slope and performance may relate to changes in neural density (i.e., node activity).

References

Arndt, J., & Reder, L. M. (2002). Word frequency and receiver operating characteristic curves in recognition memory: evidence for a dual-process interpretation. *J Exp Psychol Learn Mem Cogn, 28*(5), 830-842.

Bird, C. M., & Burgess, N. (2008). The hippocampus supports recognition memory for familiar words but not unfamiliar faces. *Curr Biol, 18*(24), 1932-1936.

Bird, C. M., Davies, R. A., Ward, J., & Burgess, N. (2011). Effects of pre-experimental knowledge on recognition memory. *Learn Mem, 18*(1), 11-14.

Bobb, A. J., Addington, A. M., Sidransky, E., Gornick, M. C., Lerch, J. P., Greenstein, D. K., . . . Rapoport, J. L. (2005). Support for association between ADHD and two candidate genes: NET1 and DRD1. *Am J Med Genet B Neuropsychiatr Genet, 134B*(1), 67-72.

Bobb, A. J., Castellanos, F. X., Addington, A. M., & Rapoport, J. L. (2005). Molecular genetic studies of ADHD: 1991 to 2004. *Am J Med Genet B Neuropsychiatr Genet, 132B*(1), 109-125.

Botha, H., & Carr, J. (2012). Attention and visual dysfunction in Parkinson's disease. *Parkinsonism Relat Disord, 18*(6), 742-747.

Breier, J. I., Gray, L. C., Klaas, P., Fletcher, J. M., & Foorman, B. (2002). Dissociation of sensitivity and response bias in children with attention deficit/hyperactivity disorder during central auditory masking. *Neuropsychology, 16*(1), 28-34.

Castellanos, F. X., Sonuga-Barke, E. J. S., Scheres, A., Martino, A. D., Hyde, C., & Walters, J. R. (2005). Varieties of attention-deficit/hyperactivity disorder-related intra-individual variability. *Biol Psychol, 57*, 1416-1423.

Dauer, W., & Przedborski, S. (2003). Parkinson's disease: mechanisms and models. *Neuron, 39*(6), 889-909.

Davidson, M. A. (2008). ADHD in adults: a review of the literature. *J Atten Disord, 11*(6), 628-641.

DeCarlo, L. T. (2002). Signal detection theory with finite mixture distributions: theoretical developments with applications to recognition memory. *Psychol Rev, 109*(4), 710-721.

DeCarlo, L. T. (2003). An application of signal detection theory with finite mixture distributions to source discrimination. *J Exp Psychol Learn Mem Cogn, 29*(5), 767-778.

Doty, R. L., Bromley, S. M., & Stern, M. B. (1995). Olfactory testing as an aid in the diagnosis of Parkinson's disease: development of optimal discrimination criteria. *Neurodegeneration, 4*(1), 93-97.

Dunn, J. C. (2004). Remember-know: a matter of confidence. *Psychol Rev, 111*(2), 524-542.

Dunn, J. C. (2008). The dimensionality of the remember-know task: a state-trace analysis. *Psychol Rev, 115*(2), 426-446.

Egan, J. P. (1958). Recognition memory and the operating characteristic. *Technical note AFCRC-TN-58-51, Hearing and communication laboratory, Indiana University*.

Ernst, M., Zametkin, A. J., Matochik, J. A., Jons, P. H., & Cohen, R. M. (1998). DOPA decarboxylase activity in attention deficit hyperactivity disorder adults. A [fluorine-18]fluorodopa positron emission tomographic study. *J Neurosci, 18*(15), 5901-5907.

Estes, W. K., & Maddox, W. T. (2002). On the processes underlying stimulus-familiarity effects in recognition of words and nonwords. *J Exp Psychol Learn Mem Cogn, 28*(6), 1003-1018.

Fearnley, J. M., & Lees, A. J. (1991). Ageing and Parkinson's disease: substantia nigra regional selectivity. *Brain, 114 (Pt 5)*, 2283-2301.

Floresco, S. B., & Grace, A. A. (2003). Gating of hippocampal-evoked activity in prefrontal cortical neurons by inputs from the mediodorsal thalamus and ventral tegmental area. *J Neurosci, 23*(9), 3930-3943.

Gilden, D. L., & Hancock, H. (2007). Response variability in attention-deficit disorders. *Psychol Sci, 18*(9), 796-802.

Glanzer, M., & Adams, J. K. (1985). The mirror effect in recognition memory. *Mem Cognit, 13*(1), 8-20.

Glanzer, M., & Adams, J. K. (1990). The mirror effect in recognition memory: Data and theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(1), 5-16.

Glanzer, M., Adams, J. K., Iverson, G. J., & Kim, K. (1993). The regularities of recognition memory. *Psychol Rev, 100*(3), 546-567.

Glanzer, M., Hilford, A., & Kim, K. (2004). Six regularities of source recognition. *J Exp Psychol Learn Mem Cogn, 30*(6), 1176-1195.

Glanzer, M., Kim, K., Hilford, A., & Adams, J. K. (1999). Slope of the receiver-operating characteristic in recognition memory. *Journal of Experimental Psychology-Learning Memory and Cognition, 25*(2), 500-513.

Gray, L. C., Breier, J. I., Foorman, B. R., & Fletcher, J. M. (2002). Continuum of impulsiveness caused by auditory masking. *Int J Pediatr Otorhinolaryngol, 66*(3), 265-272.

Gronlund, S. D., & Elam, L. E. (1994). List-length effect: Recognition accuracy and variance of underlying distributions. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 20*(6), 1355-1369.

Hagelberg, N., Martikainen, I. K., Mansikka, H., Hinkka, S., Nagren, K., Hietala, J., . . . Pertovaara, A. (2002). Dopamine D2 receptor binding in the human brain is associated with the response to painful stimulation and pain modulatory capacity. *Pain, 99*(1-2), 273-279.

Hertz, J., Krogh, A., & Palmer, R. G. (1991). Introduction to the theory of neural computation. *Addison-Wesley Publishing Company*.

Hilford, A., Glanzer, M., Kim, K., & DeCarlo, L. T. (2002). Regularities of source recognition: ROC analysis. *J Exp Psychol Gen, 131*(4), 494-510.

Hornykiewicz, O. (1998). Biochemical aspects of Parkinson's disease. *Neurology, 51*(2 Suppl 2), S2-9.

Howard, M. W., Bessette-Symons, B., Zhang, Y., & Hoyer, W. J. (2006). Aging selectively impairs recollection in recognition memory for pictures: evidence from modeling and receiver operating characteristic curves. *Psychol Aging, 21*(1), 96-106.

Hultsch, D. F., MacDonald, S. W., & Dixon, R. A. (2002). Variability in reaction time performance of younger and older adults. *J Gerontol B Psychol Sci Soc Sci, 57*(2), P101-115.

Jang, Y., Mickes, L., & Wixted, J. T. (2012). Three tests and three corrections: comment on Koen and Yonelinas (2010). *J Exp Psychol Learn Mem Cogn, 38*(2), 513-523.

Kelley, R., & Wixted, J. T. (2001). On the nature of associative information in recognition memory. *J Exp Psychol Learn Mem Cogn, 27*(3), 701-722.

Klauer, K. C., & Kellen, D. (2010). Toward a complete decision model of item and source recognition: A discrete-state approach. *Psychon Bull Rev, 17*(4), 465-478.

Koen, J. D., & Yonelinas, A. P. (2010). Memory variability is due to the contribution of recollection and familiarity, not to encoding variability. *J Exp Psychol Learn Mem Cogn, 36*(6), 1536-1542.

Koen, J. D., & Yonelinas, A. P. (2013). Still no evidence for the encoding variability hypothesis: a reply to Jang, Mickes, and Wixted (2012) and Starns, Rotello, and Ratcliff (2012). *J Exp Psychol Learn Mem Cogn, 39*(1), 304-312.

Kruschke, J. K. (2010). What to believe: Bayesian methods for data analysis. *Trends Cogn Sci, 14*(7), 293-300.

Leth-Steensen, C., Elbaz, Z. K., & Douglas, V. I. (2000). Mean response times, variability, and skew in the responding of ADHD children: a response time distributional approach. *Acta Psychologica, 104*, 167-190.

Li, S. C., & Sikström, S. (2002). Integrative neurocomputational perspectives on cognitive aging, neuromodulation, and representation. *Neuroscience & Biobehavioral Reviews, 26*, 795–808.

Lockhart, R. S., & Murdock, B. B. (1970). Memory and the theory of signal detection. *Psychological Bulletin, 74*(2), 100-109.

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

Malmberg, K. J. (2002). On the form of ROCs constructed from confidence ratings. *J Exp Psychol Learn Mem Cogn, 28*(2), 380-387.

Malmberg, K. J., Steyvers, M., Stephens, J. D., & Shiffrin, R. M. (2002). Feature frequency effects in recognition memory. *Mem Cognit, 30*(4), 607-613.

Malmberg, K. J., & Xu, J. (2006). The influence of averaging and noisy decision strategies on the recognition memory ROC. *Psychon Bull Rev, 13*(1), 99-105.

Metcalfe, J. (1993). Novelty monitoring, metacogntion, and control in a composite holographic associative recall modell: Implications for Korsakoff amnesia. *Psychological Review, 100*(1) 3-22.

Mickes, L., Johnson, E. M., & Wixted, J. T. (2010). Continuous recollection versus unitized familiarity in associative recognition. *J Exp Psychol Learn Mem Cogn, 36*(4), 843-863.

Mickes, L., Wais, P. E., & Wixted, J. T. (2009). Recollection is a continuous process: implications for dual-process theories of recognition memory. *Psychol Sci, 20*(4), 509-515.

Mickes, L., Wixted, J. T., & Wais, P. E. (2007). A direct test of the unequal-variance signal detection model of recognition memory. *Psychon Bull Rev, 14*(5), 858-865.

Mirenowicz, J., & Schultz, W. (1994). Importance of unpredictability for reward responses in primate dopamine neurons. *J Neurophysiol, 72*(2), 1024-1027.

Onyper, S. V., Zhang, Y. X., & Howard, M. W. (2010). Some-or-none recollection: Evidence from item and source memory. *J Exp Psychol Gen, 139*(2), 341-364.

Pertovaara, A., Martikainen, I. K., Hagelberg, N., Mansikka, H., Nagren, K., Hietala, J., & Scheinin, H. (2004). Striatal dopamine D2/D3 receptor availability correlates with individual response characteristics to pain. *Eur J Neurosci, 20*(6), 1587-1592.

Prull, M. W., Dawes, L. L., Martin, A. M., 3rd, Rosenberg, H. F., & Light, L. L. (2006). Recollection and familiarity in recognition memory: adult age differences and neuropsychological test correlates. *Psychol Aging, 21*(1), 107-118.

Qin, J., Raye, C. L., Johnson, M. K., & Mitchell, K. J. (2001). Source ROCs are (typically) curvilinear: comment on Yonelinas (1999). *J Exp Psychol Learn Mem Cogn, 27*(4), 1110-1115.

Quiroga, R. Q., Kreiman, G., Koch, C., & Fried, I. (2008). Sparse but not 'grandmother-cell' coding in the medial temporal lobe. *Trends Cogn Sci, 12*(3), 87-91.

Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C., & Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature, 435*(7045), 1102-1107.

Ratcliff, R., McKoon, G., & Tindall, M. (1994). Empirical generality of data from recognition memory receiver-operating characteristic functions and implications for the global memory models. *J Exp Psychol Learn Mem Cogn, 20*(4), 763-785.

Ratcliff, R., Sheu, C.-F., & Gronlund, S. D. (1992). Testing global memory models using ROC curves. *Psychological Review, 99*, 518-535.

Reder, L. M., Nhouyvanisvong, A., Schunn, C. D., Ayers, M. S., Angstadt, P., & Hiraki, K. (2000). A mechanistic account of the mirror effect for word frequency: A computational model of remember-know judgments in a continuous recognition paradigm. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 26*, 294-320.

Rolls, E. T., & Treves, A. (2011). The neuronal encoding of information in the brain. *Prog Neurobiol, 95*(3), 448-490. d

Rotello, C. M., Macmillan, N. A., & Reeder, J. A. (2004). Sum-difference theory of remembering and knowing: a two-dimensional signal-detection model. *Psychol Rev, 111*(3), 588-616.

Rouder, J. N., Pratte, M. S., & Morey, R. D. (2010). Latent mnemonic strengths are latent: a comment on Mickes, Wixted, and Wais (2007). *Psychon Bull Rev, 17*(3), 427-435.

Seamans, J. K., Floresco, S. B., & Phillips, A. G. (1998). D1 receptor modulation of hippocampal-prefrontal cortical circuits integrating spatial memory with executive functions in the rat. *Journal of Neuroscience, 18*, 1613–1621.

Seamans, J. K., & Yang, C. R. (2004). The principal features and mechanisms of dopamine modulation in the prefrontal cortex. *Prog Neurobiol, 74*(1), 1-58. doi: 10.1016/j.pneurobio.2004.05.006

Sikstrom, S. (2001). The variance theory of the mirror effect in recognition memory. *Psychon Bull Rev, 8*(3), 408-438.

Sikström, S. (2001). The variance theory for the mirror effect. *Psychonomic Bulletin & Review, 8*(3), 408-438.

Slotnick, S. D. (2010). "Remember" source memory ROCs indicate recollection is a continuous process. *Memory, 18*(1), 27-39.

Slotnick, S. D., Klein, S. A., Dodson, C. S., & Shimamura, A. P. (2000). An analysis of signal detection and threshold models of source memory. *J Exp Psychol Learn Mem Cogn, 26*(6), 1499-1517.

Solanto, M. V. (2002). Dopamine dysfunction in AD/HD: integrating clinical and basic neuroscience research. *Behav Brain Res, 130*(1-2), 65-71.

Starns, J. J., Rotello, C. M., & Ratcliff, R. (2012). Mixing strong and weak targets provides no evidence against the unequal-variance explanation of ZROC slope: a comment on Koen and Yonelinas (2010). *J Exp Psychol Learn Mem Cogn, 38*(3), 793-801.

Stenberg, G., Hellman, J., Johansson, M., & Rosen, I. (2009). Familiarity or conceptual priming: event-related potentials in name recognition. *J Cogn Neurosci, 21*(3), 447-460.

Swanson, J. M., Kinsbourne, M., Nigg, J., Lanphear, B., Stefanatos, G. A., Volkow, N., . . . Wadhwa, P. D. (2007). Etiologic subtypes of attention-deficit/hyperactivity disorder: brain imaging, molecular genetic and environmental factors and the dopamine hypothesis. *Neuropsychol Rev, 17*(1), 39-59.

Toplak, M. E., & Tannock, R. (2005). Time perception: modality and duration effects in attention-deficit/hyperactivity disorder (ADHD). *J Abnorm Child Psychol, 33*(5), 639-654.

Tripp, G., & Wickens, J. R. (2008). Research review: dopamine transfer deficit: a neurobiological theory of altered reinforcement mechanisms in ADHD. *J Child Psychol Psychiatry, 49*(7), 691-704.

Tsal, Y., Shalev, L., & Mevorach, C. (2005). The diversity of attention deficits in ADHD: the prevalence of four cognitive factors in ADHD versus controls. *J Learn Disabil, 38*(2), 142-157.

Uebel, H., Albrecht, B., Asherson, P., Borger, N. A., Butler, L., Chen, W., . . . Banaschewski, T. (2010). Performance variability, impulsivity errors and the impact of incentives as gender-independent endophenotypes for ADHD. *J Child Psychol Psychiatry, 51*(2), 210-218.

Ungless, M. A., Magill, P. J., & Bolam, J. P. (2004). Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. *Science, 303*(5666), 2040-2042.

Vincis, R., Gschwend, O., Bhaukaurally, K., Beroud, J., & Carleton, A. (2012). Dense representation of natural odorants in the mouse olfactory bulb. *Nat Neurosci, 15*(4), 537-539.

Viskontas, I. V., Quiroga, R. Q., & Fried, I. (2009). Human medial temporal lobe neurons respond preferentially to personally relevant images. *Proc Natl Acad Sci U S A, 106*(50), 21329-21334.

Volkow, N. D., Fowler, J. S., Wang, G., Ding, Y., & Gatley, S. J. (2002). Mechanism of action of methylphenidate: insights from PET imaging studies. *J Atten Disord, 6 Suppl 1*, S31-43.

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychol Rev, 114*(1), 152-176.

Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: evidence for a dual-process model. *J Exp Psychol Learn Mem Cogn, 20*(6), 1341-1354.

Yonelinas, A. P. (1999). The contribution of recollection and familiarity to recognition and source-memory judgments: a formal dual-process model and an analysis of receiver operating characteristics. *J Exp Psychol Learn Mem Cogn, 25*(6), 1415-1434.

Yonelinas, A. P. (2001). Consciousness, control, and confidence: the 3 Cs of recognition memory. *J Exp Psychol Gen, 130*(3), 361-379.

Yonelinas, A. P., Dobbins, I., Szymanski, M. D., Dhaliwal, H. S., & King, L. (1996). Signal-detection, threshold, and dual-process models of recognition memory: ROCs and conscious recollection. *Conscious Cogn, 5*(4), 418-441.

Yonelinas, A. P., & Parks, C. M. (2007). Receiver operating characteristics (ROCs) in recognition memory: a review. *Psychol Bull, 133*(5), 800-832.