



UNIVERSITÄT ZU LÜBECK

**The psychological and neural foundations of the retrieval practice
effect in young and older adulthood**

Dissertation

Submitted by

Catherine-Noémie Alexandrina Guran

Lübeck, 2020



UNIVERSITÄT ZU LÜBECK

From the Institute of Psychology I

of the University of Lübeck

Director: Prof. Nico Bunzeck

**The psychological and neurological foundations of the retrieval practice effect
and its dynamics in young and older adulthood**

Dissertation

for the Fulfillment of

Requirements

for the Doctoral Degree

of the University of Lübeck

from the Department of Natural Sciences

Submitted by:

Catherine-Noémie Alexandrina Guran

from L'Haÿ-les-Roses (France)

Lübeck 2020

First referee: Prof. Dr. rer. nat. Nico Bunzeck

Second referee:

Date of oral examination:

Approved for printing. Lübeck,

Table of Contents

Acknowledgements.....	1
Abstract	2
Zusammenfassung	5
Chapter 1 – Introduction	8
1.1 Human Long-Term Memory.....	9
1.1.1 The Atkinson-Shiffrin Model.....	9
1.1.2 Functional memory systems.....	9
1.1.3 Memory processes	11
1.2 Aging	13
1.2.1 Memory and older age	14
1.3 Retrieval Practice.....	16
1.3.1 Retrieval practice in older age.....	17
1.3.2 Temporal dynamics of the retrieval practice effect	18
1.3.3 Neural basis of the retrieval practice effect	19
1.4 Theoretical accounts of the Retrieval Practice Effect.....	20
1.4.1 Elaborative Retrieval Hypothesis.....	20
1.4.2 Retrieval as a “Fast-Route to Consolidation”	22
1.4.3 Episodic Context Account	23
1.4.4 General considerations.....	24
1.5 Goals of this dissertation	25

Chapter 2 – General Methods.....	27
2.1 Inclusion and exclusion criteria.....	27
2.2 Experimental Paradigm	28
2.2.1 Stimuli.....	30
2.3 Temperament and Character Inventory	31
2.3.1 Background.....	31
2.3.2 Administration of the TCI	33
2.4 Behavioral measurements	33
2.4.1 Experiment and analysis software.....	33
2.4.2 Participant responses	34
2.4.3 Response measures: Signal detection theory.....	34
2.4.3.1 Corrected Hit rates	34
2.4.3.2 D' (d-Prime).....	35
2.5 EEG measurements.....	36
2.5.1 Physical background	36
2.5.2 Neural oscillations	36
2.5.3 Analysis approaches: Cluster-based permutation test.....	37
2.6 fMRI measurements	38
2.6.1 fMRI rationale and physical background	38
2.6.2 Analysis approaches	40
2.6.2.1 GLM analysis.....	40
2.6.2.2 RSA (Representational Similarity Analysis)	40
Chapter 3 – Age-related decreases in the retrieval practice effect directly relate to changes in alpha-beta oscillations	42

3.1 Introduction	42
3.2 Materials and Methods	45
3.2.1 Participants.....	45
3.2.2 Experimental Design.....	45
3.2.3 EEG Data Analysis	46
3.2.3.1 Statistical Analysis of EEG data.....	47
3.3 Results.....	49
3.3.1 Behavioral results	49
3.3.1.1 Phase 1	49
3.3.1.2 Phase 2	49
3.3.1.3 Phase 3 and 4.....	50
3.3.2 EEG-Results.....	53
3.3.3 Link between behavioral and EEG results: Correlation analysis.....	55
3.4 Discussion.....	60
 Chapter 4 – Retrieval Practice Improves Recollection-Based Memory Over a Seven-Day Period in Younger and Older Adults	 65
4.1 Introduction	65
4.2 Materials and Methods	69
4.2.1 Participants and sampling	69
4.2.2 Experimental Design.....	69
4.3 Results.....	72
4.3.1 Phase 1	72
4.3.2 Phase 2	72
4.3.3 Phases 3 and 4.....	73

4.3.3.1 Main Effects.....	74
4.3.3.2 Interactions.....	75
4.3.3.2.1 Two-Way Interactions.....	75
4.3.3.2.2 Three-Way Interactions	78
4.3.4 TCI-Results	79
4.3.5 Reduced models and exploratory analyses	80
4.4 Discussion.....	81
Chapter 5 – Retrieval Practice promotes Fast Neocortical Representations in Humans	87
5.1 Introduction	87
5.2 Materials and Methods	90
5.2.1 Sample.....	90
5.2.2 Experimental Design.....	90
5.2.3 Image acquisition and preprocessing	91
5.2.4 Analysis and preprocessing of functional MRI data	92
5.2.5 ROIs	94
5.3 Results.....	96
5.3.1 Behavioral results	96
5.3.1.1 Phase 1	96
5.3.1.2 Phase 2	96
5.3.1.3 Phase 3	98
5.3.1.3.1 Main effects	98
5.3.1.3.2 Interactions.....	99
5.3.2 fMRI results	100
5.3.2.1 Flexible Factorial design – univariate analysis	100
5.3.2.2 Link between behavior and brain activation – regression analysis	106

5.3.2.3 RSA – Multivariate Mixed Model	107
5.4 Discussion.....	112
Chapter 6 – General Discussion	118
6.1. The psychological bases of the retrieval practice effect in younger and older adults	118
6.2 The neural bases of retrieval practice	121
6.2.1 RP and electrophysiological time-frequency mechanisms	121
6.2.2. Brain activation markers and activation patterns in RP	123
6.3 Implications for explanatory theories of retrieval practice	125
6.4 Future directions	131
6.5 Strengths and limitations.....	136
6.6 Conclusion	139
List of Tables	140
List of Figures	141
Glossary	143
References	144

Acknowledgements

Writing your dissertation is an exhausting, taxing, and tremendously rewarding time. While this is my achievement, I could not have done it with the help of all my co-authors, and the wonderful staff at the Institute of Psychology I, as well as everyone who helped me collect data, and my participants. My gratitude goes out to everyone involved in this project, and those people that kept me sane during it.

Thanks to my mother, my family, my dog, my best friend. Without you I would have not gotten here, been much less happy, and not made it through.

Thanks go to my friends, old and new, that have cheered me on and up during the process: Niël, Bronwyn, Patrick, Sebi, Daniel, Niki, Alex K., Gareth, Sara, and Anette; and those that were right by my side in the office, sharing in impromptu rollercoasters of emotion: Ricardo, Alex S., Le Vy, Nick, Niclas, Troby, Anke.

Thanks go to my role models that showed me the amazing world of natural sciences: Peter Lustig, my uncle Matei, and Mayim Bialik.

Thanks go to the Institute for Psychology I, Uni Lübeck, and especially Christa, Tanja, and Anne, without which this dissertation would not have come about.

Thank you Nico, for being my supervisor, with an open ear for all concerns and always pushing me to be better. I hope to make you proud. Thank you to all the rest of the Bunzeck AG: Pau, Davina, Tineke, Alex, Ricardo, Mushfa, and Anke. Wherever my endeavors, academic or otherwise, take me, I will think back fondly to the shared times.

Finally: thank you, reader, for reading my dissertation. Sweat and tears went into its making. May it be informative to you.

Abstract

The retrieval practice effect (RPE) – or “testing effect” – describes how retrieving information is a more effective way to memorize materials than re-encoding – or restudying. In this dissertation, the RP in younger and older adults, as well as electrophysiological and neural markers of the RPE were investigated.

Memory declines in older age, therefore identifying techniques that can reduce age-related memory performance impairments is particularly important, especially under the light of current demographic changes. RP is of importance on the one hand, because it is such a learning technique that can improve memory performance and accuracy, and on the other hand because studying RP increases our understanding of human memory.

In chapter 1, a brief outline of the human memory, as well as the existing findings around the RPE, is given. Three existing explanatory accounts of RP are described. Furthermore, I outline the questions that will be tackled throughout the thesis. Chapter 2 describes the general methods, namely the experimental set-up and paradigm, the methodologies (electroencephalography and functional magnetic resonance imaging) we used, and the rationale for our analyses.

In the first study (chapter 3) we used electroencephalography to investigate the neural oscillations elicited by RP in younger (18-30 years of age) and older (50+ years of age) adults. RP increased memory performance in both age groups, but to a smaller degree in older adults. On the electrophysiological level, RP led to a decrease in alpha-beta power, which is related to higher task demands and difficulty and semantic processing. Furthermore, older participants exhibited greater alpha-beta power decreases than younger participants, possibly related to the task being increasingly difficult for older adults. Importantly, the amount of RP-

specific alpha-beta power decrease was related to the individual's RP benefit in older adults only, with shallower decreases being linked to stronger RPEs, giving a direct correlate between neural signature and behavior. Alpha-beta power decrease, a marker for semantic memory, was found to play a crucial role in the RPE.

In the second study (chapter 4), we investigated the stability of RP over a 7-day period in younger and older adults, as well as the respective contributions of RP to recollection versus familiarity. RP specifically enhances recollection rates significantly more than familiarity scores, and its beneficial effects remain unchanged between immediate final recall, compared with 7-day retention interval, for both older and younger adults. These results illustrate an involvement of episodic memory that happens fast and is stable over at least a week-long period in both younger and older adults.

Based on our previous findings of fast memory improvements from RP, in the third and final study (chapter 5), we investigated predictions made by the "Fast-Route to Consolidation" account of retrieval practice (Antony et al., 2017) using functional magnetic resonance imaging. This theory views RP as leading to fast consolidation processes, which lead to enhanced neocortical integration of information. Our results showed that RP quickly enhances memory accuracy, and that this RP benefit is linked to increased activity (as measured through BOLD) in brain regions involved in consolidation, as well as semantic processing, such as the mPFC, anterior cingulate cortex, and parts of the anterior and superior temporal cortex. Our findings lend support to the FRC as well as other explanatory accounts of RP.

In chapter 6, I summarize all findings in relation to the general hypotheses of this thesis, and discuss their implications for the explanatory accounts outlined in chapter 1. Through the three studies in this dissertation, we found evidence for the involvement of semantic and

episodic memory, as well as effects attributable to encoding, consolidation, and retrieval processes respectively. The findings support different previous theories on RP, leading me to propose an integrative model of RP accounts. Finally, I offer a brief discussion of the strengths and limitations of this thesis and further questions that need to be addressed.

Zusammenfassung

Der Abruf von Informationen führt zu sogenanntem Abrufbasiertem Lernen (ABL) – auch als Testing Effekt bekannt – und ist effektiver zum Lernen neuer Materialien, als erneutes Enkodieren. In dieser Dissertation habe ich mich dem Einfluss von höherem Erwachsenenalter auf ABL, sowie den elektrophysiologischen und neuronalen Grundlagen dieses Phänomens gewidmet. Gedächtnisleistung nimmt im Alter ab: die Identifizierung von Lerntechniken, die altersbedingte Leistungsabfälle reduzieren können, ist – besonders im Lichte des demographischen Wandels – sehr wichtig. ABL ist auf Grund seiner Förderung der Gedächtnisperformanz interessant, aber auch, da die Untersuchung dieses Effektes unser Verständnis des menschlichen Gedächtnisses verbessert.

In Kapitel 1 wird ein kurzer Überblick über das menschliche Gedächtnis gegeben, sowie die aktuelle Befundlage zu ABL zusammengefasst. Drei Erklärungsansätze abrufbasierten Lernens werden ebenfalls beschrieben. Des Weiteren behandle ich die übergeordneten Hypothesen, mit denen ich mich in dieser Dissertation befasse. Kapitel 2 beschreibt die allgemeinen Methoden, insbesondere das experimentelle Paradigma, die neurokognitiven Methoden (Elektroenzephalographie und funktionelle Magnetresonanztomographie), sowie die Grundkonzepte unserer Analysen.

In der ersten Studie (Kapitel 3) haben wir mit Hilfe von Elektroenzephalographie neuronale Oszillationen betrachtet, die von ABL in jungen (ca. 18-30 Jahre) und älteren Erwachsenen (ab 50 Jahren) hervorgerufen werden. ABL erhöhte die Gedächtnisleistung in beiden Altersgruppen, jedoch etwas weniger in älteren Probanden. ABL führte zu einer Reduktion von alpha-beta power, welche mit höheren Anforderungen der Aufgabe, sowie semantischen Prozessen zusammenhängt. Zudem hatten ältere Probanden einen größeren Abfall der alpha-

beta Power in Zusammenhang mit ABL, möglicherweise durch die als höher empfundene Anforderung der Aufgabe für ältere Teilnehmer. Zu betonen ist, dass das Ausmaß der Gedächtnissteigerung durch ABL mit weniger starken alpha-beta Reduzierung in älteren Probanden zusammenhängt. Diese Reduzierung, die für semantische Gedächtnisprozesse spricht, spielt eine Rolle im ABL.

In der zweiten Studie (Kapitel 4) untersuchten wir die Stabilität abrufbasierten Lernens über einen Zeitraum von 7 Tagen, ebenfalls in jungen und älteren Erwachsenen. Auch haben wir uns mit den Beiträgen von ABL zu Erinnerungs- und Familiaritätsratings befasst. ABL steigert Erinnerungsraten deutlich mehr als Familiaritätsratings, und die positiven Effekte abrufbasierten Lernens bleiben, sowohl für junge als auch ältere Probanden, über einen Zeitraum von 7 Tagen konstant. Diese Ergebnisse verdeutlichen, dass episodisches Gedächtnis im ABL involviert ist, und ABL Prozesse schnell, sowie über den Zeitraum einer Woche in jungen und älteren Erwachsenen stabil sind.

Basierend auf den Ergebnissen der ersten beiden Studien, haben wir in der dritten und letzten Studie (Kapitel 5) mit Hilfe von funktioneller Magnetresonanztomographie einige Hypothesen der "Fast-Route to Consolidation" Theorie zu ABL getestet (Antony et al., 2017). Diese Theorie geht davon aus, dass schnelle Konsolidierungsprozesse durch ABL begünstigt werden, und zu verstärkter neokortikaler Integration von Informationen führt. Unsere Ergebnisse sprechen dafür, dass ABL zu einer schnellen Steigerung der Erinnerungsleistung für getestete Items führt. Zudem hängt das Ausmaß abrufbasierten Lernens mit neuronaler Aktivität (BOLD-Signal) im mPFC, dem anterioren Gyrus Cinguli, und Teilen des superioren Gyrus Temporalis und Temporalpols, zusammen; Regionen, die in Gedächtniskonsolidierung und semantischen

Gedächtnisprozessen involviert sind. Unsere Befunde unterstützen die FRC Theorie, sowie andere Erklärungsansätze zum ABL.

In Kapitel 6 fasse ich unsere Ergebnisse, mit Blick auf die allgemeinen Hypothesen, zusammen und diskutiere ihre Bedeutung für die in Kapitel 1 beschriebenen theoretischen Erklärungsansätze zu abrufbasiertem Lernen. Auf Basis der drei durchgeführten Studien, kann geschlussfolgert werden, dass sowohl semantisches, als auch episodisches Gedächtnis, sowie Prozesse von Enkodierung, Konsolidierung, und Abruf im ABL eine Rolle spielen. Die Ergebnisse stützen verschiedene theoretische Erklärungsansätze zu ABL, und werden letztendlich zu einem übergreifenden Modell von ABL kombiniert. Zum Schluss diskutiere ich kurz die Stärken und Schwächen dieser Dissertation und gebe einen Ausblick auf noch offene Fragestellungen.

Chapter 1 – Introduction

“(...) every day in a life fills the whole life with expectation and memory (...)” (C.S. Lewis, *Out of the Silent Planet*, Chapter 12, p. 75f)

Memory is a crucial part of the human condition: memory shapes our expectations of the present and future, and guides us in choosing courses of action suited to our environment. It has been described as the ability to “benefit from experience” (Tulving, in Gazzaniga, p.727, 2000). Human memory has been the subject of scientific thought and investigation at least since the advent of ancient Greek philosophy (Socrates suggested the advent of the written word would endanger human memory, Nehamas & Woodruff, 1995). In modern psychology, the study of human memory remains a central interest of research to this day.

Our general ability for “memory” can serve many different functions, and these functions are fulfilled by distinct neural networks, pathways, and mechanisms (Gabrieli, 1998). Influential memory models have posited distinct memory systems, which will be discussed in the following sections. Different learning techniques can be employed to commit information to memory: in particular, this thesis investigates the retrieval practice effect (RPE), which states that retrieval of information from long-term memory leads to an enhancement of memory for this information. To illustrate it with an example: when learning a poem, or a mathematical formula, one approach is to re-read it (re-encode it) multiple times until it is memorized. Alternatively, in retrieval practice, the information (poem, formula, etc.) is encoded once, and then retrieved from memory. The active retrieval from memory enhances the memory performance for the poem or formula more than repeated re-encoding (Roediger III & Karpicke, 2006). As memory performance decreases with age, and the current demographical

changes make loss of everyday function due to impaired memory performance an increasingly important problem (World Health Organization, 2015), retrieval practice (RP) is especially interesting as a learning technique that might improve memory performance in older adulthood. The aim of this dissertation is to increase our understanding of the behavioral facets of RP, and further our understanding of the neural bases of the RPE, with a focus on aging adults, in three empirical studies.

1.1 Human Long-Term Memory

1.1.1 The Atkinson-Shiffrin Model

Atkinson and Shiffrin proposed a memory model which differentiates between different depths of memory (1968, 2016), as well as an overarching (cognitive) control mechanism. Sensory input first enters the “sensory register” or sensory memory system, where a memory trace is available only briefly and is still very volatile. From there, a trace might enter the “short-term store” or working memory, which only receives and stores selected inputs. The working memory storage is also susceptible to loss of information, i.e., forgetting, unless traces are practiced – or rehearsed – through top-down control mechanisms. Finally, memory traces from the short-term store may be more permanently stored in the “long-term store”, or long-term memory. The focus of this dissertation lies on human long-term memory.

1.1.2 Functional memory systems

A highly influential proposal by Tulving (1985, 1995) posits a set of independent long-term memory systems. This model has been further adapted to differentiate between two general types of long-term memory (Schacter & Tulving, 1994; Squire, 1992): declarative and non-declarative memory. Non-declarative memory encompasses memory processes that are not

conscious, and are retrieved through actions, not primarily through verbalization, also coined procedural memory, such as knowing how to ride a bike. Another form of non-declarative memory is conceptual priming (Markowitsch, 2003), whereby concepts stored in LTM are unconsciously reactivated and this can shape following responses. Declarative memory describes information in long-term memory that mainly needs to be verbalized in order to be retrieved, such as memory for facts or events. Within declarative memory – which is the focus of this dissertation – lie further subdivisions, importantly, semantic and episodic memory (Tulving, 1985). Semantic memory encompasses the ability to represent facts about the world, without remembering the specific context in which they were learned. An example could be knowing that Mount Everest is 8848 meters high, without remembering when one first learned this information. Episodic memory includes memory for particular events, with the ability to place them in their temporal and/or spatial context, such as the memory of one's first day at university or the last movie one saw with a friend. The focus of this dissertation lies on episodic, as well as semantic memory.

Brain areas that are prominently involved in declarative memory include the medial temporal lobe (Squire & Zola-Morgan, 1991), especially the Hippocampus and surrounding areas, the prefrontal cortex (Buckner et al., 1999; Moscovitch & Winocur, 2002; Petrides et al., 1993), and the parietal cortex (Cabeza et al., 2008; Vilberg & Rugg, 2008). While the hippocampus is a key component for episodic memory (Squire & Wixted, 2011), it seems to be less relevant for the acquisition of semantic memory contents (Tulving & Markowitsch, 1998; Vargha-Khadem et al., 1997). Depending on the sensory modality of the memory trace, sensory cortices can be involved as well (Desimone et al., 1995; Harris et al., 2001; Ji & Wilson, 2007;

Weinberger, 2004), however, these will not be the focus of this work, and neither will emotional facets of memory.

1.1.3 Memory processes

Which processes lead to declarative information being stored long-term? Tulving (1995) describes how encoding, consolidation, and retrieval of memory traces are necessary for the formation, and use, of memories.

Encoding begins with information being perceived through whichever sensory modality available and attended to (i.e., visual, auditory, gustatory, olfactory, and tactile). Then, the perceived stimulus can be preliminarily stored by means of working memory (Jensen & Lisman, 2005). In terms of its neural basis, encoding new information relies on the above mentioned brain structures involved in declarative memory, such as the hippocampus and prefrontal cortex (PFC). Importantly, the (left) hippocampus is more involved in novelty processing during encoding, while the (left) PFC tracks category changes (Dolan & Fletcher, 1997).

Consolidation describes how a preliminarily stored information, or memory trace, is transferred or copied into more long-term memory stores (Stickgold, 2005), making it available for future retrieval. On a neural level, consolidation leads memory to become less hippocampus dependent, and more cortically dependent (Squire et al., 2015). This move of memory content from the hippocampus to the neocortex is generally viewed to take place offline, during sleep (Frankland & Bontempi, 2005; Stickgold, 2005). However, consolidation can also take place quickly and online, irrespective of sleep (Tse et al., 2007).

Retrieval is the act of recalling information from memory. Retrieval can make this information susceptible to forgetting (Anderson et al., 2000; Ciranni & Shimamura, 1999), alterations

(Bridge & Paller, 2012), and even strengthening, through so-called “reconsolidation” (Schwabe et al., 2014; Walker et al., 2003). Declarative content can be retrieved in different ways: firstly, by means of free or cued recall, memory content is retrieved through internal top-down mechanisms (Tomita et al., 1999), or external prompts, as are often used in memory research. Secondly, content can be recognized bottom-up, e.g. when presented with a stimulus, the participant can make a recognition judgment whether she or he has had encountered that stimulus already. Recognition memory processes can further be subdivided into feelings of recollection (or remembering) and familiarity (or knowing) (Yonelinas, 2002). Recollection is often described as an act of remembering including context and episodic information, while familiarity judgments only include a strong feeling of having encountered something before, without being able to give details as to the context or episode. These two systems of recognition of declarative memory are predominantly processed in the MTL, with recollection processes being more hippocampus and posterior parahippocampus-dependent, and familiarity processes more in the anterior parahippocampal cortex (Diana et al., 2007). Briefly, the different memory processes rely mainly on the above-mentioned brain structures involved in declarative memory, such as the hippocampus and PFC. An intricate network of areas relating to the medial temporal lobe (MTL) are relevant for retrieval (Rugg & Vilberg, 2012), as well as a reactivation of areas relevant in encoding processes (Nyberg et al., 2000). The medial PFC is involved in consolidation processes (Euston, 2007). A detailed overview of brain areas involved in different memory processes is beyond the scope of this dissertation. Further areas relevant to specific memory processes in view of our data and specific paradigm will be discussed in chapter 5. It should be noted that the phase of memory processing, accounts for how the memory systems discussed in chapter 1.1.2 Functional memory

systems(e.g. episodic and semantic memory) relate to one another. According to the serial – parallel – independent framework (Tulving, 1995), encoding takes places serially through the memory systems, while consolidation is taking place in parallel in each memory subsystem. Retrieval of content from one system is independent from the other systems. This dissertation evaluates the role all three memory processes – encoding, consolidation, and retrieval – play in retrieval practice.

1.2 Aging

Human life expectancy has seen a rapid increase over the last decade, but the human lifespan is still clearly limited. Towards its end, human life is usually affected by a notable period of decline in health and functioning, generally termed aging. Finding an exact age threshold that differentiates middle-aged from older adults is highly source dependent. The World Health Organization looks at effects of aging in adults from 60 years on (World Health Organization, 2015). The detrimental effects of aging mainly stem from cellular damage that accrues over time and a reduced ability of the organism to recover through means of cell replacement (Ferrucci et al., 2020).

The extent of age-related functional decreases is subject to a lot of inter- and intraindividual variance. Decreases in health and functioning do not start at the same age in different individuals, and not every domain of health and functioning is affected equally (e.g., sufferers of arthritis, a degenerative disease of the joints, become impaired in terms of mobility, while their cognitive functioning might be comparatively intact). Generally speaking however, both physical (Astrand et al., 1973; JudgeRoy et al., 1996) and mental performance are reduced with increasing age (Hedden & Gabrieli, 2004; Nyberg et al., 2012). There are synergies

between physical and mental health, meaning that physical and cognitive health are interdependent (Bamidis et al., 2014; Rosano et al., 2005). In this dissertation, the focus is on healthy, i.e., non-pathological, human aging and memory function.

Studying the effects of aging especially on memory is of very high importance. The absolute number of individuals aged over 60, and therefore affected by many age-related decreases in memory, health and functioning, is projected to reach 1.4 billion by 2030 (World Health Organization, 2015). Keeping people as functional as possible is of key importance in order to be able to cope with the increased health expenditure, social security, and economic demands old age poses.

1.2.1 Memory and older age

After the development of episodic memory throughout childhood and adolescences (Newcombe et al., 2007; Willoughby et al., 2012), memory remains quite stable during adulthood, up until the start of the sixth decade of life (Raz et al., 2005; Schaie, 1994). Thus, in older middle-aged adults (from 50 years of age) memory impairments can become noticeable, even in healthy individuals who do not exhibit pathological aging (e.g., do not suffer from neurodegenerative diseases). Research has consistently shown that older adults perform worse in memory tasks than younger adults (Craik & McDowd, 1987; Nyberg et al., 2012; Small, 2001). The start of this decline coincides with the decline of memory-relevant brain areas (Raz et al., 2005). Temporal and frontal areas, which are closely tied to memory processes, are particularly affected by age (Hedden & Gabrieli, 2004). In terms of general activation patterns in the brain, older participants generally show less hemispheric specialization, resulting in a relative dedifferentiation of the brain (Cabeza, 2002; Reuter-Lorenz & Park, 2010). Dedifferentiation is a general marker of neurophysiological integrity

(Lindenberger & Baltes, 1994), often associated with the loss of functional specialization in the brain (Reuter-Lorenz & Park, 2010). On the electrophysiological level, healthy older adults have been shown to have less theta power (Cummins & Finnigan, 2007), which is related to memory processes, as well as reduced beta (Gola et al., 2012) in relation to attentional processes, and reduced alpha power during memory retrieval (Karrasch et al., 2004; Sebastián et al., 2011).

However, different memory systems do not suffer equally from aging: while perceptual memory and priming stay relatively stable in healthy aging, aspects of working and declarative memory face a more serious decline (Gabrieli, 1996). Within declarative memory, semantic memory stays quite stable (St-Laurent et al., 2011), whereas episodic and autobiographical memory are more strongly affected (Spencer & Raz, 1995). Recollection judgments are more affected by normal cognitive aging than familiarity judgments (Anderson et al., 2008; Koen & Yonelinas, 2014).

Ageing affects encoding, consolidation, and retrieval. While encoding largely relies on the same networks in older and younger adults, activity in inferior PFC and hippocampus is increased bilaterally for successfully encoded items in older adults (Morcom et al., 2003). Especially when the encoding tasks rely on semantic memory and elaboration, older adults show similar neural activations as younger adults (Logan et al., 2002). In terms of consolidation, age negatively impacts sleep-related memory consolidation processes for motor skills (Fogel et al., 2014; Spencer et al., 2007) as well as memory, linked to reductions in slow-wave sleep (Backhaus et al., 2007; Mander et al., 2013). For retrieval, it has been shown that older participants have decreased activity in the left PFC and caudal temporal regions for encoding, and right PFC and parietal regions (Cabeza et al., 1997). It is important

to note, that not just losses of brain activation are noted with age; as a highly dynamic and plastic organ, the brain makes use of maintenance and compensatory mechanisms to counteract loss of function (Cabeza et al., 2018), often resulting in dedifferentiated brain activation patterns (Reuter-Lorenz & Park, 2010).

The neural changes in encoding, consolidation, and retrieval have as a consequence reduced memory performance in old age. Behaviorally speaking, successful retrieval of memory is impaired in older age. However, as impairments in each memory process can lead to reduced memory performance, it is difficult to pinpoint the cause to one specific process. Unsuccessful encoding, consolidation, and retrieval can each lead to poorer memory performance. However, based on the neural changes, all three processes seem to be affected by age. In this dissertation, age-related differences in the RPE will be investigated.

1.3 Retrieval Practice

Restudy is the most commonly employed technique for students to commit content to memory (Karpicke et al., 2009). However, research of over a century has shown retrieval practice to be more beneficial to the learning and retention of information compared with re-encoding or restudy (Abbott, 1909; Karpicke & Roediger, 2008; Spitzer, 1939; for a review see Rowland, 2014). In retrieval practice information is retrieved, instead of re-encoded, to enhance memory performance. The retrieval practice, or “testing”, effect has received continuous interest in the past two decades and is behaviorally well documented. Typically, RP paradigms consist of (1) a familiarization, or initial encoding, phase, in which memory material is encoded, (2) a practice phase, in which the material is restudied either via re-encoding or via retrieval, and (3) a final recall or recognition memory test to assess the

difference between learning techniques employed in the second phase. The recall of information through the retrieval tasks in the restudy phase then usually leads to better memory performance in comparison to simple re-encoding learning technique.

A substantial amount of evidence illustrates that retrieval practice improves memory for different types of stimulus material, such as word lists and scientific texts (see Rowland, 2014), and visual picture material (Herweg et al., 2018). Furthermore, RP manifests itself irrespective of recall modality, for both passive (recognition) and active recall (Rowland, 2014). RP benefits memory even more than elaborate encoding techniques, such as mind-mapping (Karpicke & Blunt, 2011). Even children (Fritz et al., 2007) and patients with traumatic brain injury or multiple sclerosis benefit from it (Sumowski et al., 2010, 2014). Interestingly, RP is also effective for novel stimuli: stimuli that have never been encountered before are remembered better if shown in a retrieval practice context, compared with a restudy/encoding context (Chan et al., 2006; Herweg et al., 2018). Finally, RP also selectively improves recollection compared with familiarity rates (Chan & McDermott, 2007; Verkoeijen et al., 2011), which furthers the assumption that episodic like processes might be involved in the RPE. However, a few aspects of the retrieval practice effect are still in contention. These are outlined below.

1.3.1 Retrieval practice in older age

Due to the age-related changes in neural integrity and cognitive functioning outlined above, it is possible that the RP benefits decrease. However, as older adults generally perform worse in memory tasks (Hedden & Gabrieli, 2004; Nyberg et al., 2012), it might also be conceivable for them to benefit more from RP, as there is more “room for improvement” in older adults, or that there might be no differences compared with younger adults, if RP relies on aspects of memory that are generally preserved in healthy aging, such as semantic memory (Allen et al.,

2002; Nyberg et al., 1996; St-Laurent et al., 2011). So far, a few studies have investigated age-related changes, coming to mixed results: while some find equal RPE in older and younger adults (Balota et al., 1989; Coane, 2013; Meyer & Logan, 2013; Rabinowitz & Craik, 1986), Tse et al., (2010) found that RP depended on feedback availability in older adults. Furthermore, some studies suggest that the RPE only appears after longer retention intervals in older adults (Roediger III & Butler, 2011; Roediger III & Karpicke, 2006; Toppino & Cohen, 2009). As such, the state of the literature is inconclusive regarding age-related changes in RP. In this dissertation, I investigated the age-related differences of the RPE by combining different behavioral measures with neuroimaging and EEG, in order to understand the underlying mechanisms of RP and possible divergences in the size of RP benefits in older adults that have been reported before.

1.3.2 Temporal dynamics of the retrieval practice effect

As hinted to above, some research found that the length of retention intervals is critical for the occurrence of RPEs. Not only older, but also younger adults seem to show increased RPEs with longer retention intervals between retesting and final recall (Pyc & Rawson, 2009; Rawson & Dunlosky, 2011). Some studies even suggest that the RPE is only observable after a minimum retention interval of multiple hours, to a day (Roediger III & Butler, 2011; Roediger III & Karpicke, 2006; Toppino & Cohen, 2009). However, in a study conducted in our group (Herweg et al., 2018), pictorial stimuli similar to the ones used in this dissertation showed near-immediate RP benefits. Furthermore, RP benefits seem to be quite long-lasting, spanning 16 weeks (Carpenter et al., 2009). I investigate the possible presence of immediate RP benefits throughout the dissertation and the temporal stability of RP in younger and older adults in chapter 4.

1.3.3 Neural basis of the retrieval practice effect

While the RPE has been extensively studied in behavioral settings, comparatively little is known about its neural basis. Outside of this dissertation, only few functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) studies on the RPE have been published. In terms of EEG, retrieval practice increases parietal and frontal old-new effects approximately 500-900 ms after stimulus onset, but not later (Rosburg et al., 2015). Furthermore, subsequent memory effects were found 500-700 ms after stimulus onset (and later), widely spread across the scalp in a retrieval task (Liu et al., 2017), possibly linked to episodic memory mechanisms. Importantly, Gao et al., (2016) were able to show the first clear link between ERPs and RP benefit, as in their study, ERP magnitudes at 300 to 500 ms after stimulus onset correlated with RP benefit. However, to our knowledge, there are no published studies focusing on the time-frequency domain in RP. Filling this gap in the literature could help our understanding of the underlying processes of RP, as different frequency bands of electrophysiological activity have been related to different cognitive processes, such as semantic processing for alpha-power, and theta in memory processes (Herweg et al., 2020; Klimesch, 1999).

In terms of functional magnetic resonance imaging (fMRI), performing RP increases activity in areas of working memory (parietal and frontal, Keresztes et al., 2014), and striatal activity is modulated by an interaction of restudy task format and stimulus novelty (Herweg et al., 2018), in that old stimuli elicited stronger activation if presented in a retrieval task, while new stimuli enhanced striatal activation when presented in an encoding task. Furthermore, RP increases functional connectivity between the medial prefrontal cortex (mPFC) and Hippocampus (Wing et al., 2013), as well as anterior cingulate cortex activation in a subsequent memory test

(Eriksson et al., 2011). Regarding differences between retested and restudied items at final recall, Ferreira et al. (2019) found that mPFC activity increases and hippocampal activity decreases with RP, but only after a retention interval of a minimum of two days, suggesting that RP enhances slow consolidation processes.

However, we know of no studies (published or otherwise) that have considered differences in RP processing between younger and older adults comparatively, and immediately after RP. Performing this research is important to understand the neural mechanisms behind potential age-related changes in the retrieval practice effect, as well as to expand on the existing literature by using different stimulus material and analysis methods, to answer different questions and find a link between activation and behavior. We aim to fill this gap in chapter 5.

1.4 Theoretical accounts of the Retrieval Practice Effect

In the following section, I will give a brief overview of three influential theories that try to explain how RP benefits long-term memory. Their respective qualities in light of our studies will be discussed in chapter 6.

1.4.1 Elaborative Retrieval Hypothesis

The “Elaborative Retrieval Hypothesis” (ERH, Carpenter, 2009, 2011) sees RP as a learning technique which involves mechanisms that lead to an enhancement of – or increase in – elaboration of semantic concepts in cue-target RP tasks (see also Pyc & Rawson, 2010). The hypothesis is based on the finding that weakly linked semantic cues lead to stronger RP effects for the targets than strong semantic cues (Carpenter, 2009, 2011). Conceptually, the weak cue should activate other possible cues (elaboration), which in turn function as cues for the target,

which can be used for future searches. Thereby RP increases, or elaborates, the set of viable cues for successful retrieval of the target. As an example, for the target word “shoe”, the cue word “box” is relatively weak and might elicit a larger set of following cues, such as “cardboard”, “purchase”, etc., that serve as additional cues for the activation of the target word “shoe”. A stronger cue, such as “sole”, because of its immediate connection to the target “shoe”, would not lead to an extensive elaboration of the possible cue-set. To use the example from above of learning a poem by heart, according to this account RP would lead to the activation of semantic concepts that are related to the words and verses within the poem. When trying to recollect a particular verse, the individual will then be able to use those additionally activated semantic concepts as well to recreate the verse from memory. However, it is unclear whether this elaboration occurs through narrative spreads (such as remembering episodes related to the target and cue concept, or imagining stories, see Anderson, 1983) or subconscious activation of semantically related concepts within a neural network. Narrative spreads would be slower, and explicitly experienced by participants, which has not been reported in any study of the RPE so far. The activation of semantically related concepts, or semantic elaboration in general, could be investigated through measures of neural activity. In this case, activation in areas responsible for semantic processing, or neural markers of semantic memory processes, such as alpha-power would be expected to play a role in RP.

Evidence against the ERH comes mainly from findings investigating narrative elaboration at encoding (Karpicke & Blunt, 2011; Karpicke & Smith, 2012; Lehman et al., 2014), which show that narratively elaborating on a target memory does not enhance memory performance to

the same extent as RP. Furthermore, the ERH does not make any predictions about the effectiveness of RP for non-verbal materials.

1.4.2 Retrieval as a “Fast-Route to Consolidation”

Another recent theory posits that RP functions as a fast consolidation (Antony et al., 2017). In general, memory consolidation is often viewed to be sleep dependent and rather slow (Frankland & Bontempi, 2005; Stickgold, 2005). However, RP stimuli are less susceptible to interference than restudied items (Kliegl & Bäuml, 2016) and remembered better, for longer periods of time, similar to consolidated memories (Nadel & Moscovitch, 1997). The Fast-Route to Consolidation hypothesis (FRC, Antony et al., 2017) assumes that RP enhances the integration of new information into pre-existing neocortical networks online, i.e. during awake states, while simultaneously leading to a reduction in hippocampal involvement. In our example of learning a poem by heart, RP would lead to the contents from the poem being less hippocampus dependent, and instead, more integrated within neocortical networks. RP supposedly accelerates consolidation processes to take place very quickly after the retrieval instance, thus circumventing sleep-dependent processes. Proof of concept for such quick alterations of schema-like representations in the neocortex comes from Tse et al., (2007), who trained rats in a gustatory cued recall paradigm, and observed rapid consolidation and learning in the neocortex, when information could be integrated into preexisting knowledge. As such, the FRC puts forward a few clearly testable hypotheses in terms of neural activation differences between previously retrieved and previously re-encoded items. First and foremost, RP should lead to reduced hippocampal and increased neocortex activation. Additionally, the posited process of integration into preexisting networks might lead to

stronger representational similarity in the neocortex for retrieved stimuli compared to restudied ones. These hypotheses will be addressed in chapter 5.

1.4.3 Episodic Context Account

The Episodic Context Account of retrieval practice (ECA, Karpicke et al., 2014) focusses on temporal context of memory, its reinstatement, and update, as mechanisms that are elicited by RP and enhance memory performance. Firstly, when encoding new information, episodic context information is encoded alongside the content of the information: this episodic information can be the mental state of the learner, the spatial environment, or the temporal context at time of encoding. According to Karpicke et al. (2014), the temporal context of a memory changes slowly, and is reinstated during retrieval (practice). The context information for the retrieved memory is then updated: through this reinstatement and context update, the search set of possible memories is restricted to those of a shared temporal context, and this restriction should make future retrieval more time-efficient and less error-prone. For our example of learning a poem by heart, this would mean that after initial familiarization with the poem, each subsequent retrieval attempt would lead to an attempt at reinstating the initial learning context, and in case it is different, add temporal context cues to the memory of the poem. The search-set for the poem will be restricted to the composite of temporo-contextual cues encountered during retrieval.

The ECA focusses on episodic memory as the memory system targeted by RP. The concept of temporal contexts, their reinstatement and update can be useful, and the ECA accounts for findings that show that increased spacing between retrieval instances increases the RPE. Furthermore, findings related to cue strength (see above) can also be explained by the ECA: if a cue is too weak to elicit the target on its own, the temporal context has to be reinstated to

access that memory/target (see ERH). Finally, the ECA is also useful when explaining the results of increased recollection (compared with familiarity) for items learned in RP (see Verkoeijen et al., 2011). However, the concept of temporal representations is not elaborated in the ECA (Karpicke et al., 2014), and therefore remains unprecise. Additionally, the ECA is unable to account for the results that show that novel stimuli also benefit from RP when encountered in test conditions (Chan et al., 2006; Cho et al., 2017; Pastötter & Bäuml, 2014), as no reinstatement of temporal contexts could have happened with new materials.

1.4.4 General considerations

It is important to note that the theories described above do not necessarily exclude each other, even if the authors sometimes view their theories as rivalling. Rather, they could be seen as extensions of one another: Each theory focusses on a separate memory process, and as such, their explanations could – at least to a large degree – coexist. It is conceivable that RP has effects on encoding (ERH, Carpenter, 2009, 2011), as well as consolidation (FRC, Antony et al., 2017), and retrieval (ECA, Karpicke et al., 2014). There is supportive as well as contradicting evidence for the theories outlined above, and, as will be discussed in chapter 6, this might be due to their complementary nature. Furthermore, there are a number of theories (retrieval effort/desirable difficulty hypothesis, Bjork, 1994, 1999; Pyc & Rawson, 2009; disuse hypothesis, Bjork & Bjork, 1992; Bifurcation hypothesis, Kornell et al., 2011), that while they do not exclusively deal with RP, make interesting predictions. These theories focus on the importance of effort for RP benefits: the larger the effort when retrieving an item, the more the memory trace is “strengthened”. However, the concept of strengthening is often not very clearly demarcated, and therefore of only limited use in explaining the mechanisms behind RP. As evidence for retrieval effort hypotheses can come from other sources than RP

studies (as these theories are not specific to retrieval practice), this dissertation will not provide a detailed discussion as to the evidence pertaining to these theories.

1.5 Goals of this dissertation

The aims of this dissertation are threefold: (1) to investigate the psychological, (2) the neural bases of the RPE, and (3) to identify underlying processes and mechanisms that contribute to the RPE, and to evaluate our empirical support for the existing theories on RP. In terms of the psychological bases (1), a few clear questions will be addressed: First, is RP a viable learning technique to strengthen memory traces even in older adults? Second, how (temporally) stable is the RPE, and how quickly can it strengthen memory? Third, what type of recognition memory is enhanced by RP (recollection versus familiarity) and is this the same for younger compared with older adults, who generally exhibit less recollection?

As outlined above, little is known about the neural basis of the RPE (2). Therefore, the second, but no less important aim of this dissertation was to investigate specific aspects of RP processing. In particular, the electrophysiological temporal dynamics and cognitive processes underlying retrieval versus encoding practice (chapter 3), and how RP influences memory representations, measured using fMRI at final recall (chapter 5) were investigated. Furthermore, we were interested in possible neural markers that can serve as predictors for an individual's benefit from RP, especially in old age, when episodic memory typically declines (chapters 3 and 5).

From all these specific questions regarding behavior and neural bases, however, arises the final aim of this dissertation (3): Identifying the mechanistic processes underlying RP and evaluating the theoretical accounts of the RPE. To this end, questions are: Which processes of

memory formation (i.e., encoding, consolidation, retrieval) play a role in the RPE, and how does RP affect them? How do our empirical results lend support to the different RPE theories outlined in this chapter? These questions will be addressed in the final chapter (6) of this dissertation.

Chapter 2 – General Methods

In this dissertation I performed three studies, all of which were approved by the local ethics committee of the University of Lübeck (Ratzeburger Allee 160, Haus 2, 23538 Lübeck). All participants were informed orally and in writing and gave written consent. The studies are described in the chapters 3, 4 and 5 of this dissertation. Chapter 3 is in large parts identical, or similar, to our publication entitled “Age-related decreases in the retrieval practice effect directly relate to changes in alpha-beta oscillations”, co-authored by Nora Herweg and my thesis supervisor, Nico Bunzeck. Chapter 4 is in large parts identical, or similar, to our publication “Retrieval Practice Improves Recollection-Based Memory Over a Seven-Day Period in Younger and Older Adults”, co-authored by Jovana Lehmann-Grube and my thesis supervisor, Nico Bunzeck. Chapter 5 is as of yet unpublished, but under consideration for publication. It was co-authored by Lorena Deuker, Nikolai Axmacher, Martin Göttlich and Nico Bunzeck.

2.1 Inclusion and exclusion criteria

All participants were right-handed, fluent in German, had normal or corrected-to-normal vision, no neurological or psychiatric disorders in their medical history or at time of testing, as well as no history of radiation- or chemotherapy, and no further major health issues. For our samples of young participants, we included participants between ages 18 to 30 years, whereas for the old sample, participants had to be at least 50 years old. We did not set an upper limit for age in this sample. Our threshold for the older sample group was based on research by Raz et al. (2005), which suggests that cognitive aging, especially in relation to memory, starts in

the sixth decade of life. As my dissertation focusses on healthy aging, the Montreal Cognitive Assessment was used to screen for mild cognitive impairment (MoCA, Nasreddine et al., 2005). The MoCA was preferred over the Mini Mental State Examination (Folstein et al., 1975; Tombaugh & McIntyre, 1992), as it is more sensitive to mild cognitive impairment (Hoops et al., 2009). Participants who scored < 22 points were excluded from the experiment. This threshold was established in non-clinical subject populations (Freitas et al., 2013).

2.2 Experimental Paradigm

To investigate the RPE, we adapted the paradigm from Herweg et al., (2018). The following section describes the general paradigm we used in all three studies, and this chapter 2.2 is adapted from our publications (Guran et al., 2019; Guran et al., 2020). Small alterations were made between studies which are described in the chapters pertaining to each study.

In Phase 1 (see **Figure 2.1**), participants were familiarized with 160 outdoor and indoor images (80 each), by means of a target detection task: The target stimuli were presented initially (one indoor, one outdoor) for 12s. Subsequently, the 160 images for familiarization were presented three times each, for 1s in pseudorandom order intermixed with 9% of target trials (i.e., 48 target and 480 non-target trials). Each image was followed by an inter-stimulus interval of 1.5s (white fixation cross on grey background). Participants had 2s to respond to the target stimuli with a button press and had the opportunity to pause every 96 trials. Target stimuli were not shown again outside of Phase 1.

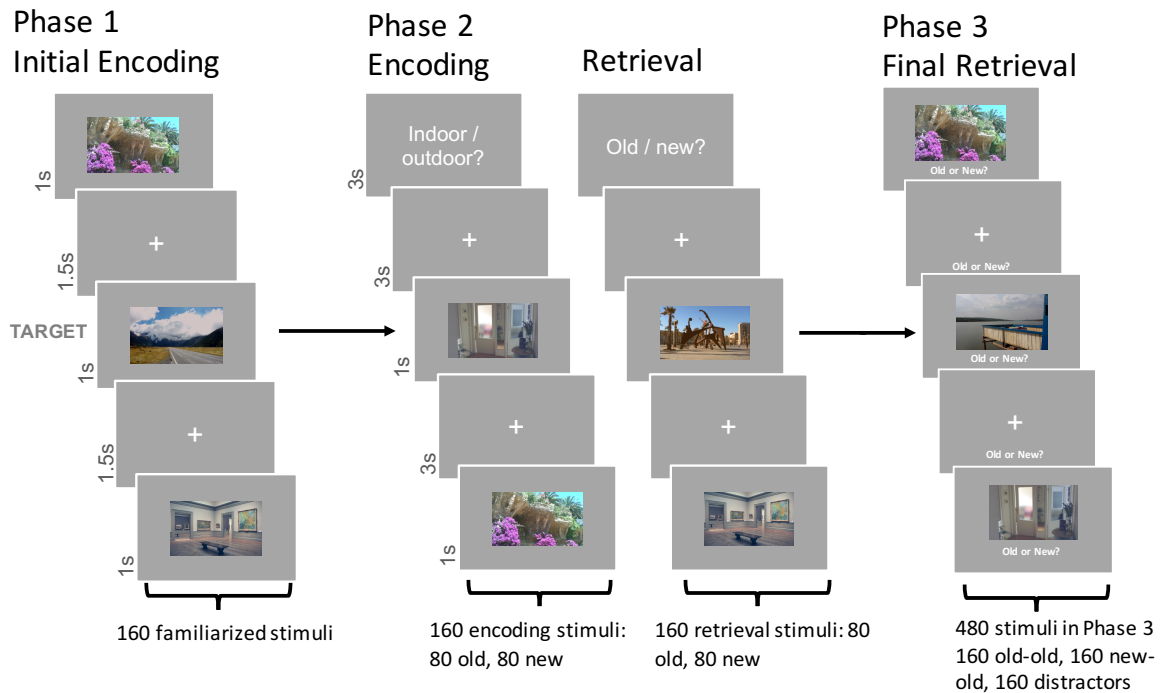


Figure 2.1: Experimental paradigm. In Phase 1, Participants were familiarized with 160 images. In Phase 2, participants saw all 160 previously shown images (old) as well as 160 new images in either one of two learning contexts: the encoding, or the retrieval task. In Phase 3, participants gave final recognition memory judgments, in different formats depending on experimental set-up. Modified from Herweg et al. (2018).

For Phase 2, participants had to perform two different, randomly alternating tasks, while viewing 160 new stimuli randomly intermixed with the 160 familiarized stimuli (within-subject). The tasks were designed to induce an encoding and a retrieval context for half of the new and familiarized stimuli each. By means of button presses, participants gave simple indoor/outdoor categorization judgments in the encoding, and old/new responses in the retrieval task. The combination of the two factors task and stimulus novelty (ENC/RET and OLD/NEW) resulted in a 2 x 2 repeated measures design with 80 stimuli per condition. Task blocks were eight trials long, each block containing four OLD and four NEW stimuli in random order. An instruction screen informed participants about the upcoming task prior to the start of each block. Images were presented for 1s with an inter-stimulus interval of 3s (fixation

cross). Participants gave their response within 2.8s using their right index and middle finger. Response-button mappings were counterbalanced across participants. Participants could make a self-paced pause every 64 trials.

In Phase 3, participants performed a surprise recognition task after the end of the second phase. The 320 previously encountered stimuli, from Phase 2, were intermixed with 160 unseen distractor images (i.e., 160 stimuli from the encoding context, 160 stimuli from the retrieval context and 160 unseen distractors). Each image was presented for 1s with either a visual analogue scale (ranging from “certainly old” to “certainly new”; chapter 3), or four response options in German (Remember – Know – Unsure – New; chapters 4 and 5) below the image. Participants were instructed orally and in writing about the meaning of each response option. They were asked to choose “Remember” when they recognized a picture and could recollect specific thoughts or associations linked to the study episode (Recollection). They were asked to choose “Know” when they recognized the picture but were not able to recall specific associations related to the study episode (Familiarity). “Unsure” was to be pressed when they did not know whether a picture was old or new, and “New” when they had not seen the picture before. The order of the response options depended on key mapping in Phase 2. The inter-stimulus interval was 3s (fixation cross and analogue scale) during which participants could still give their response using either a mouse device or a keyboard, depending on response type. Participants could make a self-paced pause every 60 trials.

2.2.1 Stimuli

High-quality, royalty-free, creative commons scene pictures were selected via the online search engines Google and Wikimedia Commons. We looked for neutral indoor and outdoor pictures of similar complexity. To not invoke emotions or strong associations, none of the

pictures showed humans or animals as their focus (e.g., a picture of an elephant in the savannah would not be chosen, while a picture of the seaside, with far-away birds in the sky, could be chosen, as those were not key features of the picture).

Images were randomly assigned to the different phases and conditions for each participant. To control for effects of illumination, mean luminance on each color channel (R,G,B) was set to 127 (scale from 0 to 288) and images were presented on a grey background of equal luminance. Prior to each phase of the experiment, participants completed a brief training session. Images used during the training phase were different from those used during testing.

2.3 Temperament and Character Inventory

2.3.1 Background

In study 1 and 2 (chapters 3 and 4 respectively), participants filled in the Temperament and Character Inventory (TCI, Cloninger, 1987) at the end of the experiment, after the final recall session. The TCI is applicable to both healthy and clinical samples, and is founded on a model of human personality that includes neurological, psychological as well as social factors. It views an individual's personality as a combination of temperament (related to procedural learning and so-called "habit systems") and character (related to propositional learning and "cognitive sets") (Garcia et al., 2017). The temperament subscales, importantly novelty seeking, have been shown to have a genetic foundation. I was especially interested in the temperament scale novelty seeking, a habit system characterized as "incentive activation". Individuals with high novelty seeking are supposed to be prone to exploring, being bored easily, enthusiasm, curiosity and impulsivity (Cloninger, 1987). The scale is further subdivided into four subscales which are:

- Exploratory Excitability versus Stoic Rigidity: this subscale is closely related to “Sensation Seeking” (McCourt et al., 1993) as it captures inclination to favor new and exciting activities or concepts.
- Impulsiveness versus Reflection: high scores on this scale describe a tendency to be impulsive and have a dramatic temperament.
- Extravagance versus Reserve: high scorers tend to not heed their resources carefully, such as money, time, or energy.
- Disorderliness versus Regimentation: this subscale describes the tendency to show anger and disregard routines.

Novelty seeking and its subscales tend to decrease with age (Cloninger, 1987, p. 89). In terms of its foundation in neurobiology, Cloninger et al. (1994) predicted novelty seeking to be linked to dopaminergic postsynaptic sensitivity and networks in the ventral tegmental area (VTA) and striatum, and a relationship between novelty seeking and substantia nigra/VTA reward processing has indeed been found (Krebs et al., 2009). Novelty seeking is linked to Dopamine D4 receptor gene polymorphisms (DRD4; Schinka et al., 2002) as well as to the Dopamine transporter gene DAT1 (Kazantseva et al., 2009). So far, most studies linking memory and Novelty Seeking come from animal models: novelty seeking in rats is linked to the integrity of the noradrenergic system (Sara et al., 1995), and is unrelated to weaning duration (Fraga et al., 2014). In humans, novelty seeking has been linked to drug consumption and abuse (Bardo et al., 1996; Wills et al., 1994) as well as novelty processing in the SN/VTA (Krebs et al., 2009). Finally, novelty seeking is negatively correlated with dopamine receptors in the human midbrain (Zald et al., 2008). Dopamine is also linked to memory functioning (Apitz & Bunzeck,

2013; Chowdhury et al., 2012; Eckart et al., 2014), and therefore, we used novelty seeking as a proxy of the dopaminergic system in chapter 4 in our investigations of the RPE.

2.3.2 Administration of the TCI

The questionnaire consists of 240 statements about the self in a “yes” or “no” format. It was given in computerized form, and, according to the manual, participants did not have a time limit to finish the questionnaire, and were told that questions should be answered based on a gut feeling and not dwelled upon to excess.

2.4 Behavioral measurements

2.4.1 Experiment and analysis software

I used MATLAB version 2016b and Psychophysics Toolbox (Version 3, Brainard, 1997) to make changes to pre-existing experimental scripts and to program the experiment. The experiment was displayed on a Windows computer in the behavioral labs in the Center for Brain, Behavior and Metabolism (CBBM, Marie-Curie Straße, 23562 Lübeck). Data were prepared with MATLAB version 2016b (chapter 3) and 2019a (chapters 4 and 5), analyzed in SPSS, version 23 (chapter 3) and 25 (chapters 4 and 5). To analyze EEG data, I used additional MATLAB toolboxes: EEGLab (Delorme & Makeig, 2004) and ADJUST (Mognon et al., 2011) for the preprocessing of EEG data, and FieldTrip (Oostenveld et al., 2011) for the analysis. In case of MRI data, I used SPM12 (The Wellcome Trust Centre for Neuroimaging, University College London), as well as R (1.1.463, R Core Team 2017, 2013), and FreeSurfer (Fischl, 2012) for brain segmentation for one of the analyses. JASP (version 0.9.2, Love et al., 2019) was used for all Bayesian statistics. Information on preprocessing of EEG and fMRI data can be found in

chapters 3 and 5 respectively. Plots were made in MATLAB with the help of the gramm-toolbox (Morel, 2018).

2.4.2 Participant responses

We recorded participants' button presses as well as response times, to be able to examine both reaction times and accuracy. The main focus of the analyses was on (memory) accuracy, while reaction times were occasionally used to assess baseline differences between age groups and tasks in Phases 1 and 2. Reaction times (RT) were not used to assess effects on memory, as participants were generally instructed to favor responding accurately over trying to respond quickly. Furthermore, we had no clear predictions as to how RP could influence RT: It is possible, that RT would decrease due to facilitated memory recall, on the other hand, RP could also lead to slower, but more accurate, memory retrieval (Reed, 1973). In future studies, it might be interesting to look at the effect of RP on RTs.

2.4.3 Response measures: Signal detection theory

We calculated corrected hit rates and d' (read: d-Prime) values for our analyses, following Signal Detection Theory (Stanislaw & Todorov, 1999), as outlined below.

2.4.3.1 Corrected Hit rates

When analyzing (memory) accuracy, simple *hit rates* can be problematic. Hit rates are calculated by dividing the number of hits by the number of targets, thus giving a probability or percentage measure indicating how often a target was identified. In a memory study, this translates to the number of old items that are correctly identified as old, divided by the total number of old items. The hit rate can, however, be biased: if a participant were to identify all presented items as old – even new ones – their hit rate would be 1, or 100%. However, their

performance would be non-discriminatory. Their *false alarm rate* (FA) would also be 1, as this is calculated as the number of incorrectly classified new items divided by the number of all new items. The *corrected hit rate* is calculated by simply subtracting the FA rate from the hit rate (Stanislaw & Todorov, 1999). Corrected hit rates vary between -1 and 1, the latter indicating perfect memory accuracy in the experiment, the former consistently incorrect responses. In memory research, corrected hit rates in the vicinity of 0 indicate little to no learning taking place.

2.4.3.2 D' (*d-Prime*)

Another measure of memory accuracy is d' . It is a z-standardization of corrected hit rates. This is achieved by using the inverse Phi distribution, with which probabilities (as calculated in hit- and FA rates) can be converted to z-scores. The z-standardized FA rate is then subtracted from the z-hit rate, analogous to the calculation of corrected hit rates. However, as the inverse phi of 0 (meaning no hits or false alarms) and 1 (meaning all hits/false alarms) is $-\infty$ and ∞ , respectively, 0.5 is added to the number of hits or false alarms, and 1 to the number of total signal trials (see Stanislaw & Todorov, 1999).

I preferentially used d' over corrected hit rates, as the z-Standardization makes d' a less biased measure. However, in chapters 3 and 4, corrected hit rate results are presented alongside d' results to show the generally high overlap between those two measures. For my calculations, hits were defined as correctly identified targets in Phase 1, correctly identified old items in the retrieval task of Phase 2, and indoor images correctly identified as indoor in the encoding task of Phase 2. In Phase 3/4, hits were correctly identified previously encountered items.

2.5 EEG measurements

2.5.1 Physical background

In the brain, information is transmitted via electrical signals from one cell to the other. This signal transmission creates a measurable electric signal on the participants scalp. More specifically, viewing the participant's skull in a simplified manner as a sphere with uniform density and electric resistance, the transmission of information via electric signaling causes electric dipoles to form (Homma et al., 1987), leading to measureable changes in electric potential on the sphere's surface, i.e., on the participant's scalp. These dipoles (like magnets) are very small differences in electric potential (voltage) from one side of the skull to the opposite. In EEG, electrodes are placed on the scalp of the participant to directly measure changes of electric potential in comparison to a reference electrode, which is positioned at a part of the skull that is less relevant to brain activity. Changes in electric potential occur very fast, so that the temporal resolution is only limited by the sampling rate of the EEG recording system that is used (see Jackson & Bolger, 2014; Luck, 2014). The voltage changes arise through local field potentials in the brain, and EEG signals are thought to arise mainly from large pyramidal neurons of the cortex, which are oriented perpendicularly to the scalp (Olejniczak, 2006). However, due to the gross approximation of the human brain within the skull as a sphere of uniform electric resistance, the spatial resolution of this methodology is quite poor.

2.5.2 Neural oscillations

The research in the first study (chapter 3) is focused on the temporal patterns of electrical activity in the brain. In particular, we were interested in changes in neural oscillations related

to memory processes of the RPE. Neural oscillations have first been described with the advent of EEG by Hans Berger in 1924 (Haas, 2003): they are rhythmic, cyclical – oscillatory – brain activity as measured through potential changes. Different frequency bands, indicating the number of cycles of potential change per second, are thought to be related to different brain functions. For our study, we focused on the alpha, beta, and theta frequency bands. Alpha power (oscillatory activity between 8 and 12 Hertz) has been related to attentional processes (e.g., Klimesch et al., 1998) and, importantly, semantic memory (Klimesch, 1996, 1999, 2012). Beta power is linked to long-range synchronization (Benchenane et al., 2011) which could be crucial for the consolidation processes that are presumably involved in RP (see above). Lastly, theta power has been linked to human memory processes numerous times (Herweg et al., 2020). In terms of time-frequency findings of memory processes, consolidation-related findings stem mainly from sleep-related offline processes (Diekelmann & Born, 2010; Marshall et al., 2006) and are therefore not of further relevance to this dissertation. Encoding and retrieval, however, have been linked to theta power increases (Hsieh & Ranganath, 2014; Klimesch et al., 2001; Osipova et al., 2006). Alpha oscillations have also been related to encoding and retrieval (Klimesch, 1999, see above).

2.5.3 Analysis approaches: Cluster-based permutation test

For the analysis of our EEG data, we performed cluster-based permutation tests (CBP, Maris & Oostenveld, 2007). The spatiotemporal and spectral nature of EEG time-frequency data poses the issue of multiple comparisons: to identify where (i.e., at which sensors), when, and in which frequency bands an effect occurs, the number of statistical tests that need to be performed will be very large. CBP has the benefit of reducing concerns arising from multiple comparisons, and, as it is a non-parametric procedure, reducing the amount of constraints on

the data and their distribution (Maris & Oostenveld, 2007). CBP can be used for both the analysis of evoked potentials as well as frequency domains, the latter being the focus of chapter 3.

The rationale of CBP centers on comparing the observed test statistic (e.g., a T-value quantifying the difference between two groups) with a Monte Carlo approximation of the (random) distribution of the data. The data from two (or more) conditions (e.g., age groups) are pooled and sampled randomly into two (or more) random samples, as would be predicted by the null hypothesis (i.e., not taking condition into account). This procedure gets repeated many times (at least 800, Pernet et al., 2015), and each t-value is saved. Finally, the initially observed t-value investigating the alternative hypothesis is compared to the distribution of t-values approximated by the Monte Carlo method: an observed t-value below a statistical threshold (e.g., an alpha-level of 0.05) indicates that the test parameters (such as power values) are different between the different conditions or groups. A detailed description of the preprocessing and analysis details is given in chapter 3.2.

2.6 fMRI measurements

2.6.1 fMRI rationale and physical background

Functional magnetic resonance imaging (fMRI) has become a gold standard for the investigation of the location of specific cognitive functions in the brain (Poldrack et al., 2011, p.1). Instead of measuring electric potentials, as seen above for EEG, fMRI signals are based on metabolic responses in the brain. Conceptually, fMRI measures metabolic demands across the brain, operationalized through blood oxygenation levels: brain areas that are actively involved in a specific cognitive task have increased metabolic demands and will need more

energy, and hence, more oxygen-rich blood, than areas that are not involved in those cognitive tasks (Poldrack et al., 2011, p.70f). Oxygenated and deoxygenated blood have different physical, and in particular magnetic, properties (Ogawa et al., 1990), which are utilized for creating images via fMRI. Specifically, participants are in a homogenous magnetic field (B_0) inside the scanner bore, which leads to many molecules in their bodies (e.g., molecules in the blood) being oriented along the field lines. To extract information on the brain's structure or function, short magnetic pulses (B_1) perpendicular to the homogenous field B_0 are transmitted by the scanner, briefly orienting molecules (especially in the blood) along the magnetic pulse's field. When the pulse is over, the molecules realign themselves to the homogenous scanner field, and this realignment of molecules (relaxation) creates small changes in the magnetic field that can be measured. In this way, a ratio between oxygenated and deoxygenated blood can be measured (BOLD response, Logothetis, 2003). fMRI can also be used to measure the concentration of certain molecules or elements in specific parts of the brain, or to illustrate different tissue types within the brain, because of the differences in magnetic properties between different molecules, or tissues.

In terms of temporal resolution, the metabolic processes captured by fMRI are very slow. First of all, before more oxygenated blood reaches the active parts of the brain, there is a deficiency of oxygen (Buxton, 2001). The peak of oxygenated blood reaching activated brain areas seems to occur at around 5 seconds after the onset of activation (Logothetis et al., 2001), at least in humans, and is therefore used when fitting a hemodynamic response function (HRF) to the fMRI signal. In terms of its relation to neuronal activity, fMRI reflects both local field potentials (input processing) and multi-unit activity (Bandettini & Ungerleider, 2001), slightly favouring multi-unit activity (Logothetis, 2003). The BOLD signal can potentially reflect various processes

of neurovascular coupling (Logothetis & Pfeuffer, 2004), as well as both pre- and postsynaptic activity (Arthurs & Boniface, 2002; Logothetis, 2008). Overall, fMRI can be used to investigate brain areas involved in a cognitive function reliably, however, fMRI data have limited resolution for the temporal dynamics of cognitive processes (Kim et al., 1997).

2.6.2 Analysis approaches

We chose to analyze the fMRI data in chapter 5 in two distinct ways, which are described below. The univariate general linear model (GLM) approach is based on activity (i.e., BOLD signal) in- or decreases, and the multivariate representational similarity analysis (RSA) on pattern similarity differences between different conditions.

2.6.2.1 GLM analysis

The most commonly used approach to analyze fMRI data is fitting a general linear model (GLM) onto the data, using the HRF. As described above, the metabolic dynamics of oxygenated and deoxygenated blood follow a well-researched time-course in the human brain (BOLD signal, Logothetis, 2003). The HRF is fitted on the data, and averages are calculated for each different condition, firstly within participants, and secondly across participants. The averaged data from different conditions can then be entered into classical frequentist statistics, such as ANOVAs and T-Tests, or linear regressions. A detailed description of the GLM analyses and its preprocessing steps is given in Chapter 5.2.

2.6.2.2 RSA (*Representational Similarity Analysis*)

Another approach to analyzing fMRI data focusses on activity patterns and their similarity – or dissimilarity (see Dimsdale-Zucker & Ranganath, 2018; Haxby et al., 2014; Kriegeskorte et al., 2008). Data are preprocessed slightly differently than in GLM analyses, and summarized into regions of interest (ROIs). Firstly, the activity pattern is extracted from a given ROI,

secondly the representational dissimilarity matrix is calculated as correlations between the activity patterns in all trials in one versus all trials from another condition. The data from each voxel (i.e., a three-dimensional pixel) loses its location information within the ROI. The data can then be predicted, compared, and analyzed with standard frequentist or Bayesian methods, however, based on the structure of our data, we used Mixed Multilevel Models, an extension of the GLM, for the analysis of the RSA data in chapter 5. A detailed description of the RSA analyses and the data preprocessing steps is given in Chapter 5.2.

Chapter 3 – Age-related decreases in the retrieval practice effect directly relate to changes in alpha-beta oscillations

This chapter is taken in large parts from Guran, Herweg & Bunzeck, 2019.

3.1 Introduction

A common approach to learning new material is to repeatedly encode the information at hand. However, retrieval practice (i.e. recall of previously encoded material from memory), is a more effective way to enhance long-term memory (Karpicke et al., 2009). This so-called ‘retrieval practice effect’ (RPE; or ‘testing effect’) has been shown for retention intervals of up to a week and various stimulus materials, such as vocabulary or scientific texts (Abbott, 1909; Karpicke & Blunt, 2011; Karpicke & Roediger, 2008; Spitzer, 1939). Importantly, while the RPE is well described in younger subjects, its age-related development and link to underlying neural oscillations remain unclear.

Learning and retrieval of declarative information gradually decline with age, which can be explained by specific structural brain changes to regions, such as the medial temporal lobe (MTL, i.e. hippocampus and surrounding cortex) and prefrontal cortex (PFC, Hedden & Gabrieli, 2004). Therefore, age-related differences of the RPE should also be expected. Indeed, the size or dynamics of the RPE appear to differ between young and older adults, which may relate to changes in costs and benefits of repeated remembering (Henkel, 2007). Similarly, RPEs in older adults are dependent on feedback availability during RP (Tse et al., 2010), and, as sometimes seen in younger subjects, they may only be observable after longer retention intervals (Roediger III & Butler, 2011; Roediger III & Karpicke, 2006; Toppino & Cohen, 2009,

but see Herweg et al., 2018; Meyer & Logan, 2013; Rowland & DeLosh, 2015 for immediate RPEs). In contrast, others indicate no age-related changes in RPEs (Balota et al., 1989; Meyer & Logan, 2013; Rabinowitz & Craik, 1986). For instance, the learning of unrelated word pairs benefits from retrieval relative to a study condition in both young and older adults (Coane, 2013). Together, RPEs can be observed in older adults, yet they are task dependent and may be less pronounced compared with younger subjects.

As to the underlying neural processes and associated age-related changes of the RPE, retrieval compared with restudy has been shown to engage a variety of brain areas, including the anterior hippocampus, lateral temporal cortices, medial PFC, striatum and dopaminergic midbrain (Herweg et al., 2018; van den Broek et al., 2013; Wing et al., 2013), as well as to enhance connectivity between the ventrolateral PFC and midline regions (Wing et al., 2013). These effects have been interpreted as reflecting semantic processing, selective strengthening of associations, retrieval effort, cognitive control, and motivational processes, which fit to behavioral interpretations of the RPE (Roediger III & Butler, 2011).

Electrophysiologically, a few EEG studies have focused on event-related potentials (ERPs). In a cue-target word pairs paradigm, subsequent memory effects at parietal electrodes (starting at around 500 ms after stimulus onset) were linked to RP conditions (Liu et al., 2017). Furthermore (in a different task), left parietal old-new effects (500-900 ms), indicating recollection, were stronger for tested than untested items (Rosburg et al., 2015). While both studies provide an EEG marker for the RPE, a more direct link comes from Gao et al. (2016), demonstrating that ERP magnitudes (300-500 ms after stimulus onset) in the retrieval condition correlate with RP benefits.

Apart from ERPs, neural oscillations, in particular within the theta (4-8 Hz), alpha (9-12 Hz), and beta (13-25 Hz) bands, can provide specific insights into the underlying neural mechanisms of learning and memory processes (e.g. Düzel et al., 2010; Hanslmayr & Staudigl, 2014). However, to our knowledge, no study has been published on how age-related changes in RPEs relate to specific neural oscillations. Therefore, we used a modified version of a previously established paradigm (Herweg et al., 2018) as well as EEG and time-frequency analysis (TFA). We expected RPEs in young and older adults at both time points (day of encoding [day 1] and subsequent day [day 2]), and reduced memory performance in general, as well as a reduced RPE in older adults. Despite mixed evidence at the behavioral level (see above), these hypotheses are primarily based on our own work (Herweg et al., 2018), showing an early RPE in the young, and on the proposed link between the RPE and brain regions, such as the MTL and PFC (Herweg et al., 2018; van den Broek et al., 2013; Wing et al., 2013), which typically show age-related degeneration leading to impaired behavior (Hedden & Gabrieli, 2004). These behavioral effects should be mirrored in theta, alpha and beta oscillations. While we had no clear hypothesis regarding the direction of theta and beta changes, alpha power decreases were expected in the retrieval condition due to higher task demands (e.g. Klimesch, 1999).

3.2 Materials and Methods

3.2.1 Participants

In this study, I tested 33 young (18 to 30 years) and 31 older (from 50 years on) participants. The lowest MoCA score in this sample of older adults was 25/30. From the initial sample, four participants had to be excluded due to technical difficulties, and another five due to other problems discovered during screening in the lab (left-handedness and history of psychiatric diseases). This led to a total of 27 young participants (14 male, age range: 18-29, mean age 21.4, SD=2.7), and 28 older participants (13 male, age range: 53-80, mean age 62.8, SD=7.6).

3.2.2 Experimental Design and Task

In this study, Phase 1 and 2 were performed as described in chapter 2.2 Experimental Paradigm During Phase 2, taking place approximately 30 minutes after the first phase, EEG was recorded with a BrainVision 64-electrode active EEG-system. The final recall (Phase 3) was split into two equal parts, each part containing one hundred sixty previously encountered stimuli, counterbalanced for location (indoor/outdoor), stimulus Novelty and Task in Phase 2, intermixed with 80 unseen distractor images (i.e., 80 stimuli from the encoding context, 80 stimuli from the retrieval context and 80 unseen distractors). The first part took place directly after Phase 2, and the second one on the following day (Phase 4, containing the other set of 160 familiar, and 80 novel stimuli), roughly 24 hours after Phases 1 through 3. Participants were asked to give a rating of memory confidence on a visual analogue scale, ranking from “sure old” to “sure new” for each image. The procedure can be seen in **Figure 3.1**. After completing Phase 4 of the experiment, participants filled out a computerized German version of the Temperament and Character Inventory (TCI, Cloninger, 1987).

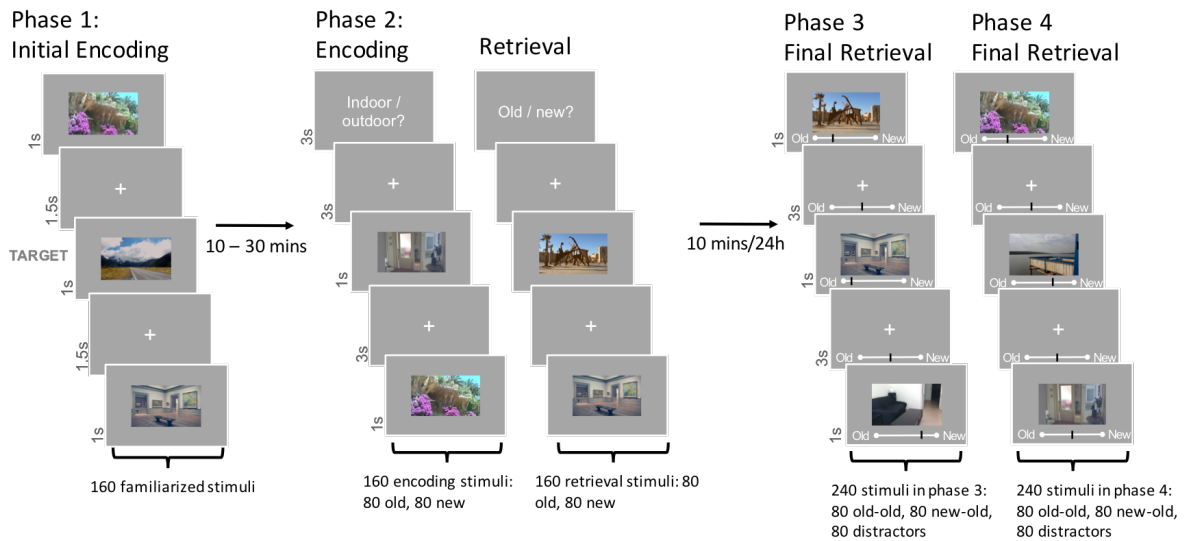


Figure 3.1: Experimental paradigm. Participants were familiarized with 160 images during the first phase (p1). In Phase 2, participants saw all 160 previously shown images (old) intermixed with 160 new images in either the encoding, or the retrieval task. In Phases 3 and 4, participants saw a randomized counterbalanced half of all previously shown images (old-old and new-old; 80 per category and phase) together with 80 new distractor images each (also 80 per category and phase). Phases 3 and 4 were performed 10 min or 24 h after Phase 2, respectively, and participants gave recognition memory strength ratings. Modified with permission from Herweg et al. (2018).

3.2.3 EEG Data Analysis

EEG activity was acquired with an EasyCap system (BrainProducts) with 64 standard active electrodes. For detecting vertical and horizontal eye movement, electrodes with indices 29 through 32 were used (VEOG/EOG). Impedances were maintained $< 10 \text{ k}\Omega$. An electrode on location TP8 served as reference, and an additional electrode placed behind the right ear as ground electrode.

The sampling rate was at 500 Hz with online high-pass (0.1 Hz) and low-pass (140 Hz) filters. EEGLAB (version 13; Delorme and Makeig, 2004) and customized MATLAB (The MathWorks; RRID:SCR_001622) tools were used for preprocessing the EEG data offline. First, data were

high-pass (1 Hz), low-pass (60 Hz), and notch-filtered (50 Hz). Second, data was inspected visually for paroxysmal artifacts. All trials were epoched (5-second trials, from 2 seconds before stimulus onset to 3 seconds after) and down-sampled to 250 Hz. The ICA was performed and components to remove were selected with help of the ADJUST toolbox (Mognon et al., 2011), which suggests components for removal, as well as visual inspection of all components. A second visual inspection for, and removal of, remaining artifacts was performed. Finally, data were re-referenced to the average reference.

Baseline correction from 1.5 to 1.4 seconds before stimulus onset and data analysis were performed using the FieldTrip toolbox (Oostenveld et al., 2011, <http://www.ru.nl/neuroimaging/fieldtrip>). Time-frequency decompositions were conducted from 4 to 30 Hz, across each epoch from 2000 ms before to 3000 ms after stimulus onset, using convolution on the single-trial time series with complex Morlet wavelets (width = 4), with steps of 4 ms in the time and 0.5 Hz in the frequency domain. For each condition, power was averaged across trials.

3.2.3.1 Statistical Analysis of EEG data

To test for differences between the power bands in the different tasks and novelty conditions, two-tailed non-parametric cluster-based permutation tests (CBP) (Maris & Oostenveld, 2007), with a family-wise error level of 0.05, were performed. All scalp electrodes were included in the test. T-tests were performed for each sample: Within-subject factors were analyzed with repeated measures T-tests, while the between-subjects factor was analyzed in an independent T-test. Clusters were formed based on adjacent significant samples ($p < 0.05$), considering only effects with three or more significant neighboring channels. To control for multiple comparisons, a Monte Carlo estimate of the permutation p-value was computed by

randomly permuting condition labels ($N = 1000$). The proportion of surrogate clusters whose cluster statistics (sum of t-values) exceeded that of clusters observed with original condition labels constitutes the corrected p-value. Non-parametric CBP tests, as described above, were conducted on the time-frequency data from 0 ms to 2000 ms after stimulus onset, and 4 Hz to 30 Hz. While CBP tests are a very robust method to analyze EEG data, they are unsuitable for ANOVA analyses. To assess interactions we used T-Tests on the averages over conditions. To correlate EEG findings with behavioral data, we extracted power values from peak electrodes within a cluster (visual inspection), which were then Pearson correlated with d' difference values. To compare correlations between age groups, we used the CoCor software (Diedenhofen & Musch, 2015). All figures were created in MATLAB, with the help of the EEGLAB toolbox (Delorme & Makeig, 2004), and in case of the bar plots, the gramm toolbox (Morel, 2018).

3.3 Results

3.3.1 Behavioral results

3.3.1.1 Phase 1

Participants' accuracy in the first phase, which consisted of a target detection task, was similarly high in both age groups ($d'_{\text{young}} (M \pm SD) = 5.3 \pm 0.36$, $d'_{\text{older}} = 5.47 \pm 0.3$), and a T-Test yielded no significant difference between the groups ($p > 0.2$).

3.3.1.2 Phase 2

The number of correct and incorrect responses per condition can be seen in **Table 3.1**. In Phase 2, data were analyzed in a 2 x 2 ANOVA, with task (ENC vs. RET) as a within-subjects factor and age group as between-subjects factor.

Table 3.1: Number of averaged correct and incorrect responses per condition, for Phase 2, encoding (indoor vs. outdoor) and retrieval task (old vs. new).

Age group	Encoding		Retrieval			
	Correct	Incorrect	Old	Old	New	New
			Remembered	Forgotten	Correct Rejections	False Alarms
young	151.04	6.78	49.48	28.96	66.44	12.63
older	154.54	3.36	57.25	20.96	58.07	20.64

There was a main effect of task for response accuracy (d' , ENC/RET; $F_{1,53} = 541.19$, $p < 0.001$, $\eta_p^2 = 0.911$, **Table 3.2**), as participants responded more accurately in the encoding than the retrieval task. There was a significant interaction of task and age group (ENC/RET vs. young/older: $F_{1,53} = 4.7$, $p = 0.036$, partial $\eta^2 = 0.081$, see **Table 3.2**). The difference in accuracy between retrieval and encoding task was significantly smaller in young participants than in older ($t_{1,53} = 2.16$, $p = 0.036$). There was no main effect of age group ($p > 0.39$). As d' for the

retrieval condition was calculated based on responses for old and new images, these data cannot be analyzed in a 2 x 2 ANOVA.

Table 3.2: d' values for young and older participants in Phase 2. Performance was overall better in the encoding task. Older participants' performance difference between encoding and retrieval was larger.

Age	Task	Mean	SEM ^a
Young	Encoding	3.55	0.14
	Retrieval	1.44	0.13
Older	Encoding	3.9	0.12
	Retrieval	1.36	0.14

^aStandard error of the mean.

3.3.1.3 Phase 3 and 4

The number of correct and incorrect response per condition can be seen in **Table 3.3** as well as **Figure 3.2**.

Table 3.3: Number of averaged correct and incorrect responses per condition for Phase 3 and 4.

Age group	Previously encoded stimuli				Previously retrieved stimuli			
	Old in Phase 2		New in Phase 2		Old in Phase 2		New in Phase 2	
	Remembered	Forgotten	Remembered	Forgotten	Remembered	Forgotten	Remembered	Forgotten
<i>Phase 3, final recall on day 1</i>								
Young	32.85	6.89	23.04	16.56	36.26	3.56	28.3	11.3
Older	32.79	6.68	24.79	14.64	34.46	4.86	26.21	12.96
<i>Phase 4, final recall on day 2</i>								
Young	28.04	11.7	17.07	22.7	32	7.74	23.11	16.67
Older	30.82	8.82	20.6	19.29	32.5	7.36	22.6	16.93

We conducted 2 x 2 x 2 x 2 ANOVAs, with age group as a between-subjects factor, and retrieval day (Phase 3/Phase4), task (ENC/RET), and novelty (OLD/NEW) as within-subject factors for each dependent variable, namely, d' as well as memory strength ratings.

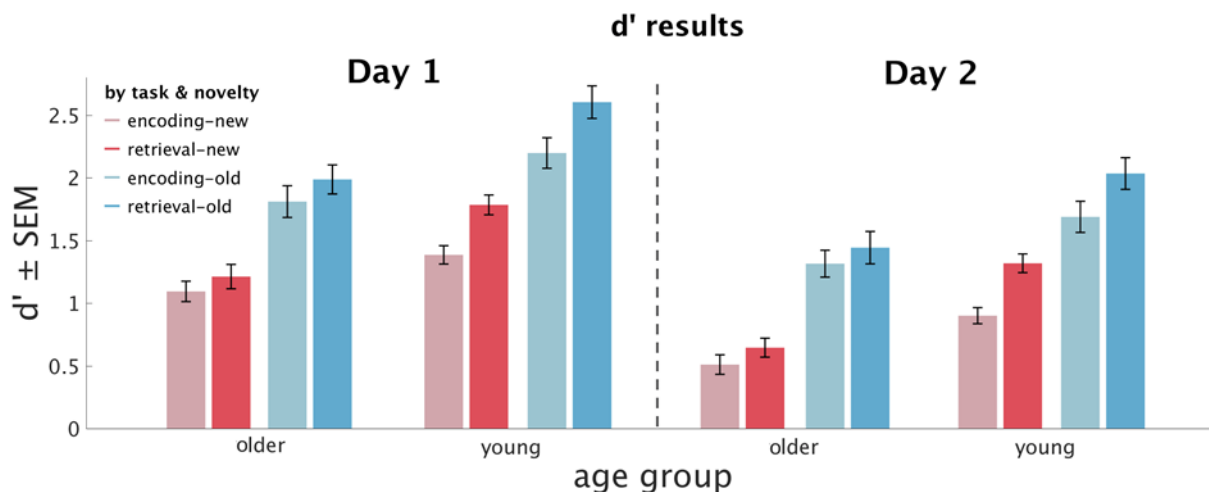


Figure 3.2: Descriptive results. Memory accuracy in Phase 3 (first experimental day) and 4 (second day), expressed in d' for older and young participants, encoded, retrieved, new, and old stimuli, respectively.

There were main effects (**Figure 3.3**) of age group (young/older, $F_{1,53} = 16.16$, $p < 0.001$, partial $\eta^2 = 0.23$), retrieval day (Phase3/Phase4, $F_{1,53} = 181.71$, $p < 0.001$, partial $\eta^2 = 0.77$), task (ENC/RET, $F_{1,53} = 83.3$, $p < 0.001$, partial $\eta^2 = 0.61$), and stimulus novelty (OLD/NEW, $F_{1,53} = 270.4$, $p < 0.001$, partial $\eta^2 = 0.84$) on memory accuracy (d'). Young participants were more accurate than older adults, participants gave more accurate memory judgments on the first experimental day, they were more accurate for old stimuli (i.e., which were already present in Phase 1, therefore old in Phase 2) than for new images, and they were more accurate for stimuli presented in the retrieval task than for those presented in the encoding task.

Importantly, there was a significant interaction between task and age group (ENC/RET x young/older: $F_{1,53} = 18.82$, $p < 0.001$, partial $\eta^2 = 0.26$): this interaction was driven by a stronger memory benefit in the retrieval task as compared to the encoding task for young versus older participants ($t_{1,53} = 4.338$, $p < 0.001$, **Figure 3.4**). In other words, the RPE was more pronounced

in the young as compared to older participants. There were no other significant interactions ($p > 0.13$).

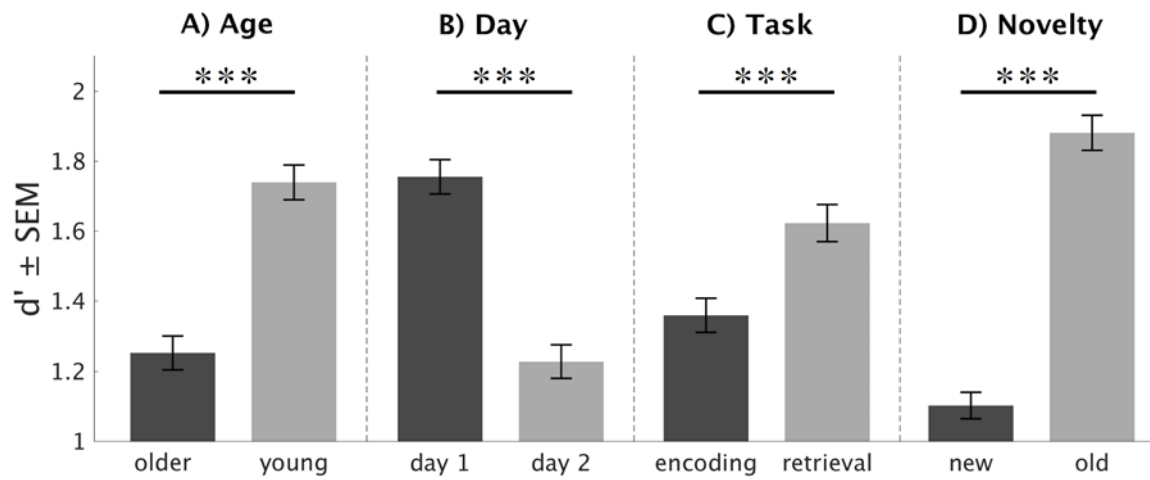


Figure 3.3: Main effects of recognition memory performance in Phases 3 and 4. Recognition memory performance (d') is depicted for age (A), task (C), and novelty (D) across Phases 3 and 4 (i.e., day 1 and 2, respectively), and for retrieval day (B). All factors (age, day, task, novelty) revealed significant main effects. $***p < 0.001$.

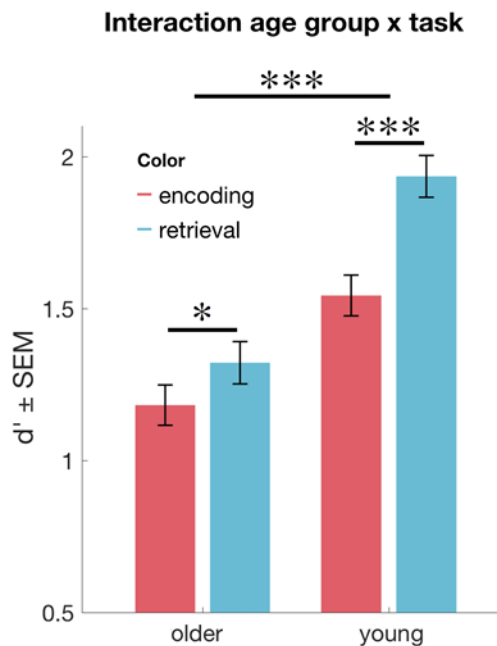


Figure 3.4: Age-dependent RPEs. Across Phases 3 and 4 (i.e., day 1 and 2, respectively), recognition memory performance was lower in older adults. Importantly, older participants showed a significant RPE, which was significantly lower compared with the younger participants. $***p < 0.001$.

The findings from the memory strength ratings replicated these results in the main effects, with lower memory strength on the second retrieval day (Phase3/Phase4: $F_{1,53} = 68.94$, $p < 0.001$, partial $\eta^2 = 0.57$), higher memory strength for stimuli from the retrieval task (ENC/RET: $F_{1,53} = 87.14$, $p < 0.001$, partial $\eta^2 = 0.62$), and higher memory strength for previously old stimuli (OLD/NEW: $F_{1,53} = 255.56$, $p < 0.001$, partial $\eta^2 = 0.83$) with one exception: age group did not have a significant influence on the memory strength ratings ($F_{1,53} = 0.42$, $p > 0.8$). The interaction between task and age ($F_{1,53} = 17.52$, $p < 0.001$, partial $\eta^2 = 0.25$) was compatible with the d' analysis and showed a significantly smaller RPE for memory strength ratings for older than for young participants ($t_{1,53} = 4.19$, $p < 0.001$). In addition to the findings in the d' ratings, there were two significant interactions in memory strength ratings: day by age group ($F_{1,53} = 4.06$, $p = 0.049$, partial $\eta^2 = 0.071$), and day by novelty ($F_{1,53} = 7.95$, $p < 0.01$, partial $\eta^2 = 0.13$). For the former, memory strength ratings dropped less from one day to another for the older participants, while the young ones showed a more pronounced decrease in memory strength ratings ($t_{53} = 2.02$, $p = 0.049$), while for the latter, memory strength decreased more strongly for novel stimuli than for old ones ($t_{54} = 2.85$, $p < 0.01$). No other effects were significant ($p > 0.05$).

3.3.2 EEG-Results

In a first step, Monte Carlo cluster-based permutation tests were run on the time-frequency data of young and older participants grouped together, from 0 to 2000 ms after picture onset, and from 4 to 30 Hz. Most effects were widespread in terms of time, significant electrodes and frequency (see **Table 3.4**): therefore, we describe power peaks after visual inspection of the significant differences between conditions. We found a main effect of task (RET vs. ENC), with decreased frontocentral alpha-beta power (10 – 20 Hz) in retrieval from 1.0 to 1.4 s, a

main effect of novelty (OLD vs. NEW), with decreased occipital and left frontal alpha power (8 – 10 Hz) for old stimuli from 0.8 to 1.1 s, a main effect of age group (young vs. older), with stronger frontocentral beta power (16 – 23 Hz) for young participants from 1.0 to 1.4 s (see **Figure 3.5, A, B and C** respectively). Importantly, there was an interaction between task and age group, with stronger frontocentral alpha-beta power (10 – 20 Hz) differences for retrieval versus encoding in older participants in comparison to young ones, from 1.0 to 1.4 s (**Figure 3.6**). To further assess these effects, we performed separate CBP tests for distinct frequency bands: alpha (9-12 Hz), lower beta (13-19 Hz), and higher beta (20-30 Hz). Again, it revealed statistically significant effects in all three frequency bands (with $p < 0.01$, $p < 0.05$, $p < 0.01$, respectively), that were similar in terms of electrode location and temporal dynamics compared with the initial analysis (4-30 Hz). Results for the analyses of separate frequency bands are not shown but will be made available upon request.

Finally, the interaction between novelty and age group was significant ($p = 0.044$, data not shown). Post hoc analysis revealed that the contrasts old versus new stimuli calculated for older participants did not reach significance, whereas the contrast for young participants only reached trend level ($p = 0.052$) suggesting a bigger power difference between old and new stimuli for young participants.

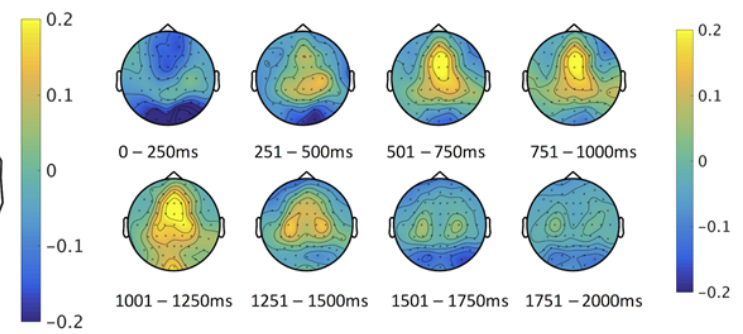
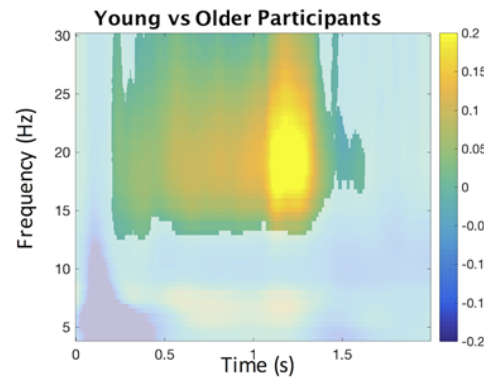
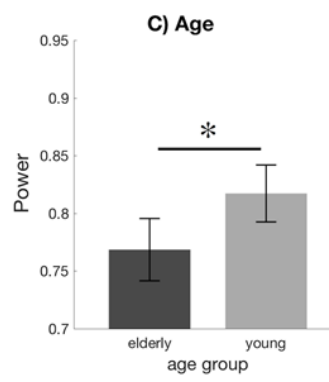
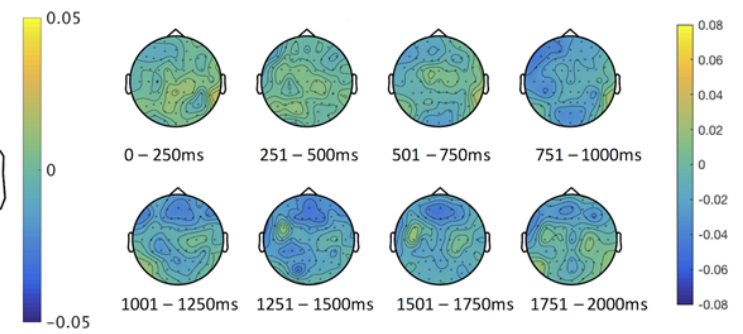
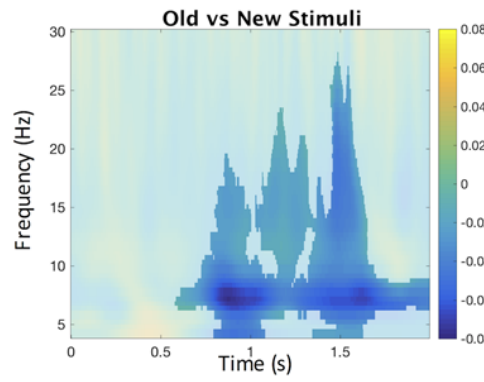
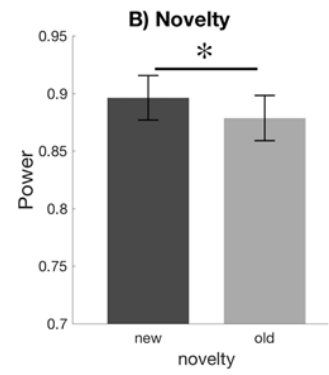
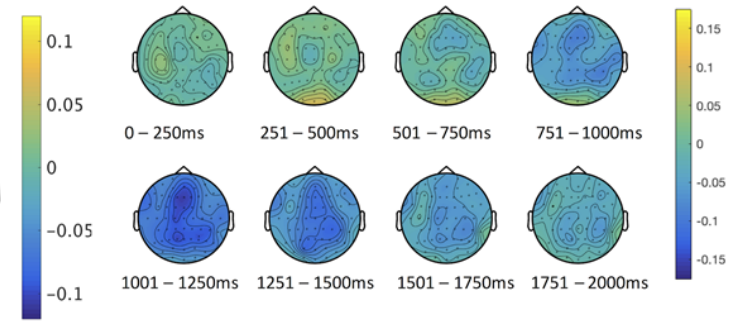
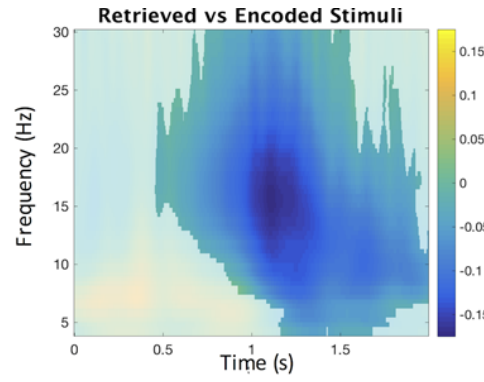
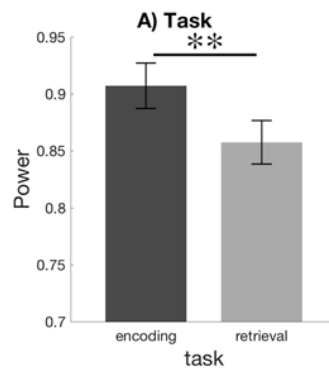
The cluster-based permutation analysis revealed no significant results for the comparisons of novelty and task, or novelty, task, and age. Contrasts for each cluster are indicated above. The significance level, duration in milliseconds, and the list of electrodes that were detected in the cluster are shown for each contrast (**Table 3.4**). **Figure 3.5** and **Figure 3.6** show the main effects and interaction between task and age group, respectively.

Table 3.4: Results of the TFA.

Cluster	Frequency range	Duration	Significant electrodes	p
Retrieval vs. encoding	4 – 30 Hz	456 – 2000 ms	All channels	Negative, 0.004
Old vs. new	4 – 28 Hz	580 – 2000 ms	All but TP8	Negative, 0.012
Young vs. older	12 – 30 Hz	200 – 1630 ms	All but T8	Positive, 0.04
Retrieval vs. encoding: young vs. older	5 – 30 Hz	960 – 2000 ms	All channels	Positive, 0.024
Old vs. New: Young vs. older	4 – 26 Hz	660 – 2000 ms	All but P8 & TP8	Negative, 0.044

3.3.3 Link between behavioral and EEG results: Correlation analysis

To further assess the relationship between task-related behavior (i.e., RPE) and neural oscillations (i.e., power changes), we extracted power values for all electrodes within the clusters, and calculated Pearson correlations between a) the power difference between encoding and retrieval task in Phase 2 and b) memory accuracy (d') difference between previously encoded versus retrieved stimuli (across both recall Phases 3 and 4). Specifically, visual inspection revealed slight differences between young and older participants in terms of peak electrodes (see **Figure 3.7A** and **below**, topoplots). For younger participants, the average power (8-25 Hz) was strongest on electrodes F2 and F4 (in accordance with visual inspection, see **Figure 3.7B**, whereas for older participants, the average power (8-25 Hz) was strongest on electrodes FC1, Cz, and F1. In order to have the same number of electrodes and comparable locations between both age groups, we chose FC1 and F1 for the analysis in older participants. For older participants, we found a significant positive correlation between the memory accuracy difference between encoding and retrieval ($d'_{ENC} - d'_{RET}$) at final recall in



(from previous page)

Figure 3.5: Time frequency results. Main effects of (A) task (retrieval vs. encoding). Participants have a higher deactivation in the retrieval task, across high alpha and low beta. B, Novelty (old vs. new). Participants have a higher deactivation for old stimuli. C, Age (young vs. older). Young participants have stronger beta power at 1–1.4 s than older participants. Raw power values are plotted in bar plots. Single topoplot across the frequency peak of the cluster. Time-frequency samples not significant on any channel within the cluster are displayed opaque. Multiple topoplots represent the time course of the cluster. Color bars represent the relative change in power from baseline. * $p < 0.05$, ** $p < 0.01$.

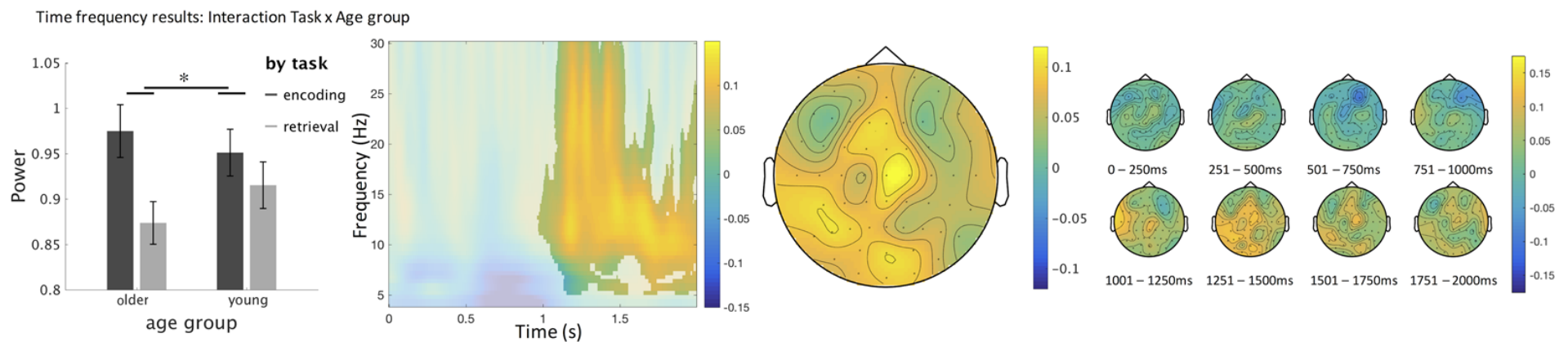


Figure 3.6: Interaction of task and age, calculated by subtracting the difference between encoding and retrieval in the young participants from the difference in the older participants, which have a stronger deactivation in retrieval than young participants. * $p < 0.05$. Raw power values are plotted in a bar graph to illustrate the interaction. Single topoplot across the visual frequency peak (8–30 Hz) of the cluster (see singleplot). Time-frequency samples not significant on any channel within the cluster are displayed opaque. Multiple topoplots represent the time course of the cluster. Color bars represent the relative change in power from baseline.

Phases 3 and 4 on one hand, and the activation difference between ENC and RET on the other hand ($r_{ENC/RET} = .41$, $p = 0.033$, **Figure 3.7A**). There was no such correlations for young participants ($p > 0.3$, **Figure 3.7B**). A direct comparison between the correlations of the older and young participants showed a significant difference (Fisher's Z : -2.12 , $p = 0.034$).

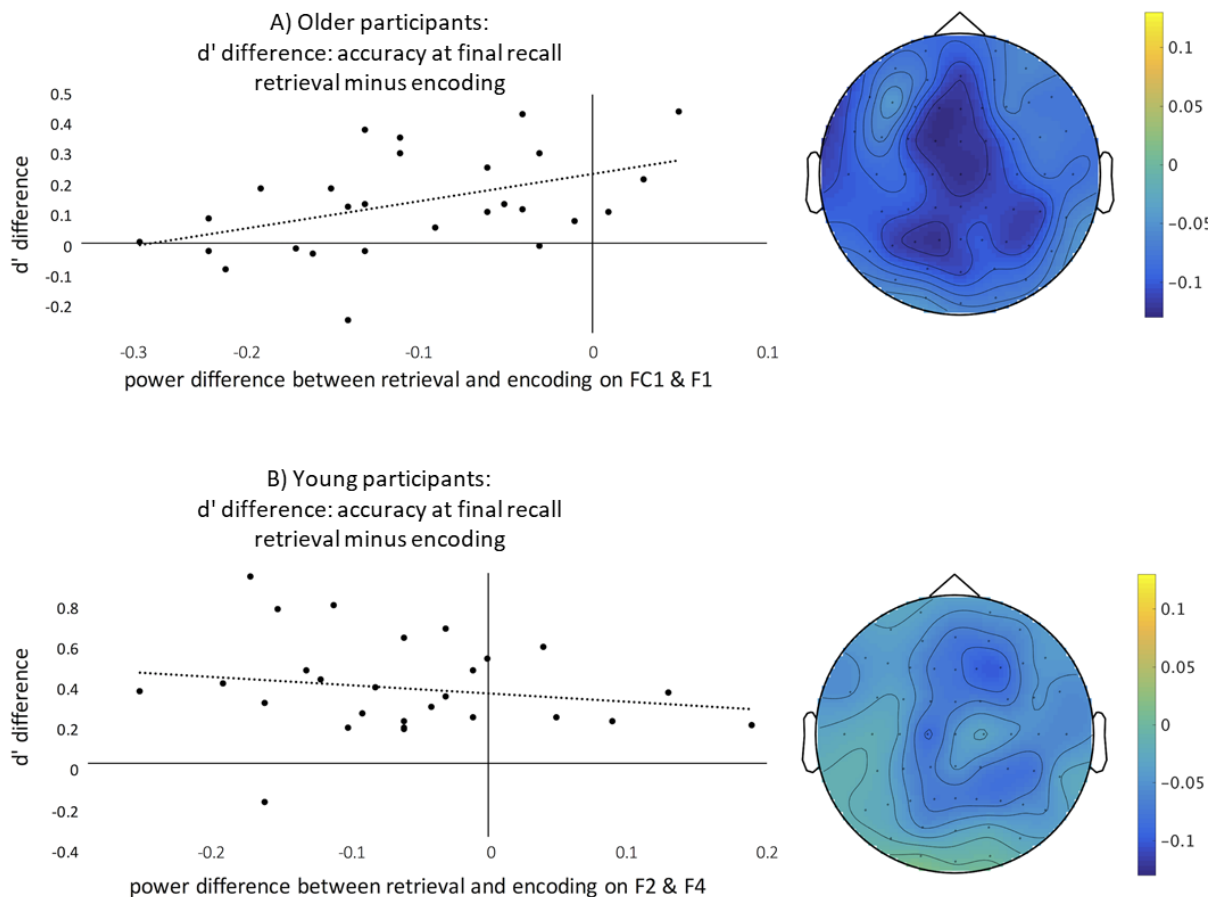


Figure 3.7: Correlation of power difference between retrieval and encoding in the central electrodes, and memory accuracy. This correlation was significant in the older, $r = 0.405$, (A) but not in the young participants, (B). Topoplots of retrieval condition minus encoding for older and young participants, respectively, over peak frequencies (8 –25 Hz). The effects are broadly centrally located, whereas the peak for the young participants is shifted frontally.

Finally, we did not perform a subsequent memory analysis (i.e., difference due to later memory; DM-analysis) since there were not enough trials in some of the four possible conditions (in particular for forgotten stimuli, see **Table 3.3**).

3.4 Discussion

Retrieval practice (RP) is a well-described strategy to improve long-term memory, but the underlying neural processes and possible age-related changes remain unclear. In this study, we can show that both young and older participants benefit from RP but the effect on subsequent long-term memory is reduced in older adults. At the electrophysiological level, retrieval (compared to encoding) was, across age groups, associated with decreased alpha/beta power, which indicates increased attentional demands and possibly deeper semantic processing (Klimesch et al., 1997; Klimesch, 2012). Importantly, the retrieval related alpha/beta decrease was more pronounced in older adults, suggesting a close relationship to age-related changes in the RPE. Indeed, correlation analyses revealed that older participants with larger RPEs show a more youth-like alpha/beta response, whereas those with small RPEs show a larger alpha/beta drop during retrieval. As such, our findings indicate a functional role and age-related changes of alpha/beta oscillations in the RPE across the life-span.

As expected, long-term memory in both young and older participants was improved by RP (**Figure 3.3C**). This is in line with long-standing evidence that testing of information leads to better memory compared with repeated study (Abbott, 1909; Karpicke & Blunt, 2011; Karpicke & Roediger, 2008; Spitzer, 1939). While this is true for various stimulus material and retention intervals, the literature is less conclusive regarding age-related changes. While some report no age effects (Balota et al., 1989; Meyer & Logan, 2013; Rabinowitz & Craik, 1986), our findings are compatible with others (e.g. Rogalski et al., 2014; Tse et al., 2010) by showing that older participants benefit from retrieval but – importantly – their RPE was less pronounced. We speculate that differences in study designs (i.e., stimulus material, retention

interval and type of retrieval) might help to explain divergent findings of age-related changes in RPE, which need to be addressed in future studies.

Overall, recognition memory was (a) lower for the older participants (**Figure 3.3A**), and (b) it declined from the early (day of encoding) to the late (one day after encoding) time point (**Figure 3.3B**). This is both compatible with previous literature on long-term memory in older adults (e.g. Hedden & Gabrieli, 2004; Nyberg et al., 2012), and the notion of how we forget over time (e.g. Oberauer & Lewandowsky, 2008; Wixted, 2004). More importantly, however, the RPE was, in both groups, observed at both time points independent from the retention interval (i.e., no interaction between task by day, **Figure 3.2**). Specifically, an RPE on day one replicates our previous observation in younger subjects (Herweg et al., 2018) and is compatible with the hypothesis of a “fast-route to consolidation”: accordingly, retrieval leads to fast consolidation by enhancing the integration of new information into preexisting neocortical networks (Antony et al., 2017). Specifically, this mechanism is supposed to rely on fast modifications of neocortical representations, which have been shown in schema-based learning paradigms (Tse et al., 2007).

Another important replication of our previous study (Herweg et al., 2018) is an RPE for those images that were novel in Phase 2 (see Chan et al., 2006; Cho et al., 2017 for a similar effect). That means that retrieval mode not only drives long-term memory for retrieved old stimuli but also other, i.e., new stimuli, that are presented in the same context. While this cannot easily be explained by the “fast-route to consolidation” hypothesis, it might be attributable to enhanced difficulty (or effort) associated with retrieval (Bjork, 1999; Rowland, 2014), or a broader search-strategy and activation of related semantic concepts (Chan et al., 2006). Accordingly, more effort may lead to deeper information processing, which is associated with

slower response times (as in our study) and increased activity in brain regions that drive long-term memory performance, such as the mesolimbic system (Boehler et al., 2011; Herweg et al., 2018; van den Broek et al., 2013).

Note that the “fast-route to consolidation” hypothesis is mainly supported by our behavioral findings: the EEG activity, on the other hand, was not acquired during the final retrieval, which would be most informative regarding the “fast-route to consolidation” framework. Finally, it should be noted that an RPE on day one is partly incompatible with other studies (Roediger III & Karpicke, 2006b), which further suggests that the RPE depends on factors such as stimulus material (verbal vs. pictures), and retrieval mode (free recall vs. recognition) (e.g. Roediger III & Butler, 2011).

At the neural level, the retrieval of information (in comparison to repeated encoding) led to power decreases in the alpha/beta band (**Figure 3.5A**). This effect was most pronounced at frontal, but also central and parietal, electrodes with a peak at 1000-1250 ms after stimulus onset. In general terms, alpha oscillations have been associated with attentional processes, and, more specifically, alpha power decreases with attentional and semantic memory demands (see e.g. Klimesch, 1999), which nicely fit the notion of enhanced task difficulty/effort during retrieval outlined above. This is further supported by the interaction between task and age (**Figure 3.6**), which was driven by stronger alpha/beta suppression in the older participants and may reflect increased cognitive effort during retrieval due to age related impairments (see below for further discussion of the age effect).

A common effect of retrieval on both alpha and beta is compatible with previous studies, which have linked both frequency bands with memory processes. In particular, alpha/beta power decreases have been observed during the encoding of verbal information, which might

reflect semantic processing and drive subsequent memory performance (for a review, see Hanslmayr et al., 2012). Interestingly, decreased alpha/beta power also relates to memory retrieval, indicating material specific memory reactivation (Hanslmayr et al., 2012). In our study, a direct comparison between encoding and retrieval task revealed lower alpha/beta power during the retrieval of information (**Figure 3.5A**). Importantly, this effect was more pronounced in older participants, suggesting age-related changes with a link to memory performance. Indeed, the correlation analysis between RPE and alpha/beta power (i.e., difference in alpha/beta between encoding versus retrieval mode, **Figure 3.7**) demonstrates that older adults with a more pronounced RPE show a more youth like alpha/beta pattern (i.e., a less pronounced power difference), whereas those with a smaller RPE show a more pronounced alpha/beta difference. While this age-related effect can be explained by compensation or dedifferentiation (see below), the lack of a correlation in the young participants is unlikely to be due to differences in variance and/or overall higher performance in this group (see e.g. **Table 3.1** and **Figure 3.4**).

From a more mechanistic point of view, both alpha and beta oscillations have been associated with several functions that are mostly complementary rather than contradictory. For instance, alpha/beta predominantly represent top-down-directed influences (Fries, 2015), whereas beta oscillations appear to represent long-range communication between brain regions, preservation of the current brain state, and endogenous (re)activation of cortical representations (Engel & Fries, 2010; Fries, 2015; Kopell et al., 2000; Spitzer & Haegens, 2017). How can age-related changes in neural activity be interpreted? The more pronounced difference in alpha/beta power for retrieval versus encoding in the older participants (**Figure 3.6**) is compatible with previous studies. Specifically, changes in neural activity in older adults,

often expressed as increased fMRI activity, have been reported before (Cabeza et al., 2002; Dolcos et al., 2002). This may either reflect compensatory mechanisms in order to preserve cognitive functioning (Barulli & Stern, 2013), or it may reflect an age-related loss of regional specificity (i.e., dedifferentiation hypothesis, Reuter-Lorenz & Park, 2010). While the compensation hypothesis might be more parsimonious in explaining our group differences, the correlation analysis is in favor of the dedifferentiation hypothesis. Indeed, the compensation hypothesis would predict that high performing older adults show more compensatory activity in order to reach the behavioral level of the younger ones. Instead, the high performing older adults showed a more youth-like EEG response. While our study was not designed to specifically disentangle both hypotheses, recent methodological advances (i.e., multivariate tests) may help to further address this issue (Morcom & Henson, 2018) as well as the relationship to possible underlying anatomical brain changes (Hedden & Gabrieli, 2004).

Together, young and older participants' long-term memory benefits from retrieval. In both groups, retrieval was associated with decreases in alpha/beta power, which indicates a direct link and might reflect attentional demands associated with increased effort. This was further supported by reduced RPEs in the older participants, which were associated with a more pronounced alpha/beta drop. As such, our findings provide evidence that older adults benefit less from retrieval practice, and they suggest that changes in alpha-beta oscillations play a direct role.

Chapter 4 – Retrieval Practice Improves Recollection-Based Memory Over a Seven-Day Period in Younger and Older Adults

This chapter is taken in large parts from Guran, Lehmann-Grube and Bunzeck, 2020.

4.1 Introduction

Repeated encoding is a common way of learning new materials, however, retrieving new material (i.e., attempting to recall material) is more effective (Karpicke et al., 2009). In terms of temporal dynamics, the RPE is considered to be stable across retention intervals (i.e., the interval between RP and final recall) of days, and weeks (Bouwmeester & Verkoeijen, 2011; Butler & Roediger III, 2007). However, the literature diverges on whether it is present immediately after RP (Roediger III & Karpicke, 2006; Toppino & Cohen, 2009), and whether its strength decreases, increases or remains unchanged with longer retention intervals. One possible explanation for these discrepancies is that previous studies used different designs, for instance, regarding how retrieval was practiced (e.g., free recall vs. recognition), and the nature of subsequent memory tests. Apart from open questions regarding the temporal stability of the RPE, possible age-related changes, as well as the underlying mechanisms, also remain unclear. Addressing these issues was the focus here.

In view of its pervasive long-lasting memory effects, RP is a potential candidate to improve memory not only in young but also older adults. This is particularly important since older age is typically associated with a decline in declarative memory functions, starting as young as 50 years of age (Crook et al., 1986). At the neural level, these declines have been linked to specific brain structures (Hedden & Gabrieli, 2004). Particularly, the medial temporal lobe (MTL),

including the hippocampus and surrounding cortex which plays a key role in recognition memory, declines with age, starting – on average – in the middle of the sixth life decade, with considerable inter-individual variation (Raz et al., 2005). Recognition memory is associated with the MTL (Diana et al., 2007), and is often investigated using the remember/know paradigm (Tulving, 1985), which assumes that recognition can either be associated with specific details or associations of the encoding episode (i.e., recollection) or take place in the absence of such associative-recollective experience (i.e., familiarity). Support for this dual-process idea (Yonelinas et al., 1996, 2010) comes from functional imaging studies that suggest different regions of the MTL to be involved in the two different memory experiences.

A few recent studies indicate that the RPE may be impaired with age: while RP has been shown to still provide memory benefits in older adults (Meyer & Logan, 2013), the benefit is smaller than in young adults (Guran, Herweg, & Bunzeck, 2019), and sometimes depends on other characteristics of the task such as feedback (Tse et al., 2010). Thus, in general, the effect appears less robust in older than in young adults, and age-related neurobiological changes might play a role in explaining the behavioral differences in RP between age groups. However, it still remains unclear whether, in the context of RP, age affects specific forms of recognition memory (i.e., recollection vs. familiarity), and whether the effects show the same temporal profiles as in young participants.

A range of (mechanistic) theories has been proposed to explain the RPE, such as the Episodic Context Account (Karpicke et al., 2014) and the Fast-Route to Consolidation Hypothesis Antony et al. (2017) (see Chapter 1.4). However, empirical evidence for either theory remains scarce since only a few studies have investigated the RPE in a source recognition memory paradigm (or investigated the RPE in the MR scanner). Existing studies indicate specific effects

of RP on recollection but not familiarity ratings (Chan & McDermott, 2007), which would be in line with the Episodic Context Account Theory. However, the temporal stability of recognition-specific RPE, as well as the rapidity with which RP benefits become apparent, and possible age-related changes remain unclear.

The personality trait “Novelty Seeking,” which is associated with exploratory activity in response to novelty, has been linked with the brain’s dopaminergic system (Benjamin et al., 1996; Ebstein et al., 1996), including the substantia nigra/ventral tegmental area (SN/VTA) (Krebs et al., 2009). This is an interesting observation since dopamine also plays an important role in memory functions (Chowdhury et al., 2012; Schott et al., 2006; Wittmann et al., 2005), and its bioavailability is known to decrease in old age, especially in the SN/VTA due to transporter and receptor loss (Bäckman et al., 2006; Fearnley & Lees, 1991). Under the assumption that Novelty Seeking is an indirect measure of the brain’s dopaminergic system, then individuals with higher Novelty Seeking traits might, at the same time, show a better functioning memory, as well as larger increases in memory due to RP.

To further investigate the RPE, especially regarding temporal stability and possible age-related changes, a recollection/familiarity approach (Tulving, 1985; Yonelinas, 2002) was used. Specifically, a group of older and younger participants were presented with a series of scene images both in an encoding and retrieval context. Their recognition memory for the presented images was tested on the same day, and 1 week after the initial learning phases using a remember/know paradigm. We expected an RPE in both young and older adults (hypothesis 1), which should be reduced in the latter (hypothesis 2). The central aim of this study was to investigate the roles of recollection and familiarity in the RPE, and we expected a remember-specific RPE (hypothesis 3). Due to our previous work (Guran et al., 2019; Herweg et al., 2018),

we also expected the RPE to be present immediately after RP and a week later (hypothesis 4), irrespective of stimulus novelty (Cho et al., 2017; Guran et al., 2019) (hypothesis 5). In addition, we explored a possible relationship between Novelty Seeking and the RPE (hypothesis 6).

4.2 Materials and Methods

4.2.1 Participants and sampling

Participants were recruited online using the Online Recruitment System for Economic Experiments (ORSEE, Greiner, 2015). A total of 49 young participants, and 46 older participants were measured. From this initial sample, nine participants had to be excluded due to a below-threshold MoCA score (one), later reported brain hemorrhages (one), technical difficulties (two), no show to the follow-up measurement (two), falling asleep during measurement (one), or not following task instructions (two). Demographic information on the remaining sample can be found in **Table 4.1**.

Table 4.1: Demographic data for the younger and older subsample.

Age group	Age		Sex		MoCA		Total
	Mean \pm SD	Range	♀	♂	Mean \pm SD	Range	
Young	24.29 \pm 3.57	18-33	23	22	--		45
Older	67.61 \pm 8.3	50-82	19	22	26.6 \pm 1,8	23-30	41

4.2.2 Experimental Design

In this study, Phases 1 and 2 were performed as described in chapter 2.2 Experimental Paradigm. Phase 3 was again split into two halves, as in the previous study described in Chapter 3 – Age-related decreases in the retrieval practice effect directly relate to changes in alpha-beta oscillations. Each half contained one hundred sixty previously encountered stimuli, counterbalanced for location (indoor/outdoor), stimulus novelty and task in Phase 2, intermixed with 80 unseen distractor images (i.e., 80 stimuli from the encoding context, 80 stimuli from the retrieval context and 80 unseen distractors). The first half (Phase 3) was completed shortly after Phase 2, the second half (with the other set of 160 familiar, and 80

novel stimuli) was performed 7 days after the participants had completed Phases 1 through 3 (Phase 4). Participants gave their answers in the form of source memory judgments, with the options “New”, “Unsure”, “Know”, and “Remember”. The experimental paradigm can be seen in **Figure 4.1**.

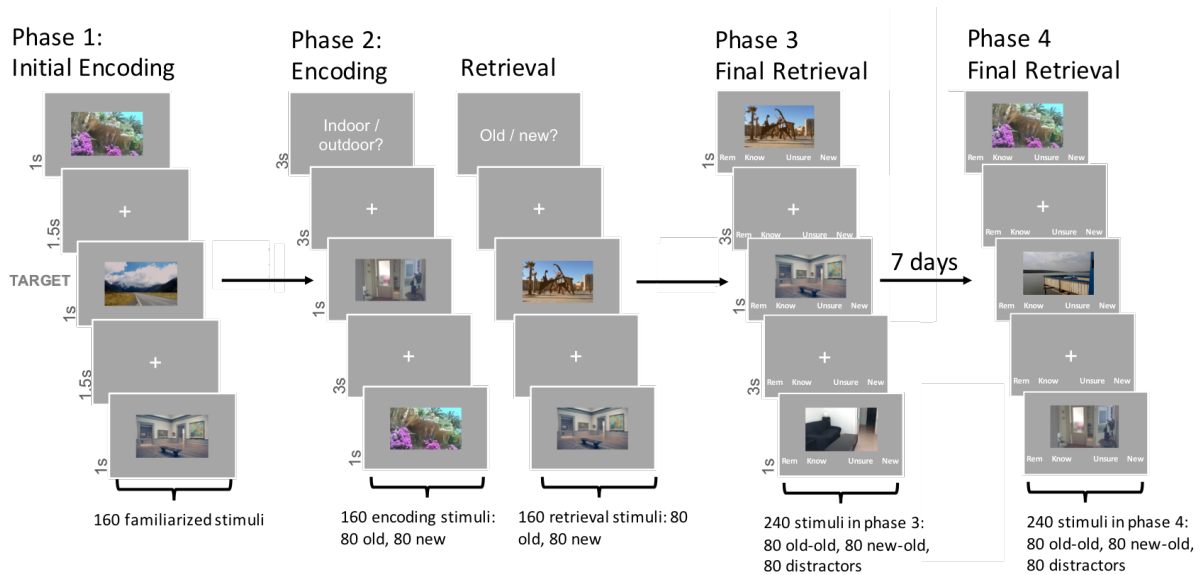


Figure 4.1: Experimental paradigm. Phase 1 familiarizes participants with 160 stimuli through a target detection task. Stimuli first shown in Phase 1 (OLD stimuli) are then shown in Phase 2, interspersed with 160 NEW stimuli, in either one of two tasks, a) the Encoding Task, in which images are categorized into indoor or outdoor, or b) the Retrieval Task, in which participants have to determine whether they have seen the stimulus already or not. In Phase 3, half of all previously encountered stimuli are shown with 80 new distractors and participants have to recall the stimuli, giving recollection/familiarity responses in form of “remember” or “know” responses. Phase 4 follows the same rationale as Phase 3, with the other half of stimuli, and a week later.

Phases 1, 2, and 3 were all performed consecutively, with small breaks (5–10 min) between them. After completing Phase 4 of the experiment, participants filled out a computerized German version of the Temperament and Character Inventory (TCI, Cloninger, 1987). The 55 TCI datasets of participants collected in the previous study (chapter 3) were available and

included in the TCI analysis in this chapter. The analysis methods of behavioral data are described in chapter 2.4.

4.3 Results

4.3.1 Phase 1

Participants' accuracy in the first phase, which consisted of a target detection task, was similarly high in both age groups, $d'_{\text{young}}(M \pm SD) = 5.36 \pm 0.33$, $d'_{\text{older}} = 5.53 \pm 0.23$. As the assumption of homogeneity of variance was violated, both for the mean (Levene-Test: $F_{\text{Mean}} = 10.93$, $p < 0.01$) and the median statistic ($F_{\text{Median}} = 6.84$, $p < 0.05$), an independent-samples T-Test with bias-corrected bootstrapping (1000 samples) was performed. The results suggest a higher performance in the older age group ($T_{84} = -2.62$, $p_{\text{bootstrapped}} = 0.008$).

4.3.2 Phase 2

In Phase 2, data were analyzed in a 2 x 2 repeated measures ANOVA, with task (ENC/RET) as a within-subjects factor, and age group as a between-subjects factor. Data residuals were normally distributed except d'_{Encoding} in the older sample, $p < 0.001$ (Kolmogorov–Smirnov Test: $D_{41} = 0.23$), and variance was homogenous (lowest $p > 0.36$). There was a main effect of task for response accuracy d' (task: $F_{1,84} = 256.53$, $p < 0.001$, partial $\eta^2 = 0.753$); a post hoc paired T-Test revealed that participants responded more accurately in the encoding than the RET ($T_{85} = 16.13$, $p < 0.001$). There was no main effect of age group, and no significant interaction of task and age group ($p > 0.8$). Since d' for the retrieval condition was calculated based on responses for old and new images, these data cannot be analyzed in a 2 x 2 x 2 repeated measures ANOVA.

Furthermore, we calculated a 2 x 2 ANOVA based on the reaction times for each task and each age group. There was a main effect of task ($F_{1,84} = 245.21$, $p < 0.001$, partial $\eta^2 = 0.745$), and of age group ($F_{1,84} = 12.74$, $p < 0.01$, partial $\eta^2 = 0.98$). There was no interaction ($p > 0.19$).

While homogeneity of variance was present (smallest p -value = 0.2), data residuals were nonnormally distributed for the young participants ($p < 0.05$), thus we used bootstrapped T-Tests. Both younger and older participants were faster in the Encoding as compared to the Retrieval task (young: $T_{44} = -9.7$, $p_{\text{bootstrapped}} < 0.01$, older: $T_{40} = -12.77$, $p_{\text{bootstrapped}} < 0.01$; $\alpha = 0.025$). Younger participants were faster than older participants in both task conditions (Encoding: $T_{84} = -3.23$, $p_{\text{bootstrapped}} < 0.01$, Retrieval: $T_{84} = -3.58$, $p_{\text{bootstrapped}} < 0.01$; $\alpha = 0.025$). The means and SDs of reaction times can be found in **Table 4.2**.

Table 4.2: Means and standard deviations of reaction times in the different tasks in Phase 2.

		Mean (msec)		SD (msec)	
		Mean (msec)	SD (msec)	Mean (msec)	SD (msec)
Age group	Young	Encoding	785	127	
		Retrieval	924	153	
	Older	Encoding	887	147	
		Retrieval	1043	153	

4.3.3 Phases 3 and 4

Addressing our hypotheses 1–5, Phases 3 and 4 were analyzed together as part of one $2 \times 2 \times 2 \times 2 \times 2$ repeated measures ANOVA, with memory (remember/know), retrieval day (Day 1/Day 7), task (ENC/RET), stimulus novelty (old/new stimuli) as within-subjects factors, and age group (young/older) as between-subjects factor. Homogeneity of variance was violated in only three of the 16 variables used for the ANOVA, smallest $p = 0.034$ (Levene-Test). Normality of residuals was violated in six out of the 32 variables (16 split by age group), smallest $p = 0.001$ (Kolmogorov–Smirnov). However, as the sample size was large ($n > 40$ per age group) and there do not seem to be suitable non-parametric analogs of a mixed $2 \times 2 \times 2 \times 2 \times 2$ ANOVA, we conducted the ANOVA as planned. We corrected for multiple tests for the effects of interest (see hypotheses 1–5). The Bonferroni-adjusted α -level for main effects and

interactions of interest in this ANOVA was $p < 0.01$. For all other effects, results were exploratory, although when correcting for all tests conducted by the 5-way ANOVA (31 in total), almost all remained significant nonetheless ($\alpha = 0.0016$).

4.3.3.1 Main Effects

Regarding our hypothesis 1 of increased memory accuracy in the Retrieval Task, there was a main effect of task ($F_{1,84} = 101.77$, $p < 0.001$, partial $\eta^2 = 0.55$), as well as main effects of memory ($F_{1,84} = 89.24$, $p < 0.001$, partial $\eta^2 = 0.52$), day ($F_{1,84} = 208.3$, $p < 0.001$, partial $\eta^2 = 0.71$), stimulus novelty ($F_{1,84} = 360.54$, $p < 0.001$, partial $\eta^2 = 0.81$), and age group ($F_{1,84} = 34.51$, $p < 0.001$, partial $\eta^2 = 0.291$). As can be seen in **Figure 4.2**, stimuli that had been presented in the RET in Phase 2 were remembered significantly better than stimuli presented in the ENC ($T_{85} = -11.18$, $p < 0.001$). Furthermore, participants showed higher remember than know rates ($T_{85} = 9.52$, $p < 0.001$), higher memory accuracy on the first day in comparison to the recall 7 days later ($T_{85} = 22.42$, $p < 0.001$), and higher memory accuracy for older stimuli, i.e., those that were shown initially in Phase 1 ($T_{85} = 21.29$, $p < 0.001$). Younger participants had higher scores as compared to older participants ($T_{84} = 3.57$, $p < 0.01$).

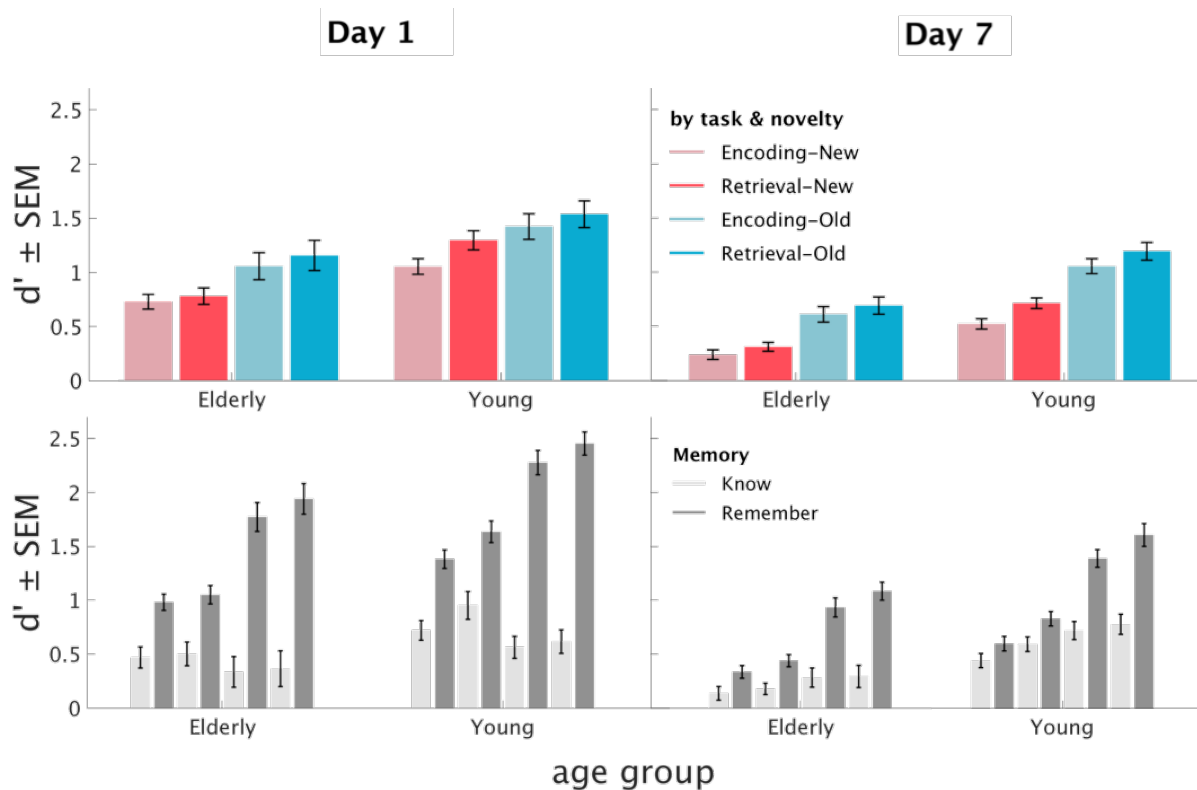


Figure 4.2: Retrieval performance for Day 1 and Day 7. D' values for all conditions: Day (separate plots), Age group (x-axis), Memory (lightness), Task (colour saturation), Novelty (colour temperature, blue = OLD).

4.3.3.2 Interactions

In total, there were five significant 2-way interactions, and three significant 3-way interactions. Figures are provided for interactions that were related to the hypotheses. T-Tests were performed to disentangle the two-way interactions, as well as for the three-way interaction, which related to one of the hypotheses.

4.3.3.2.1 Two-Way Interactions

Regarding our hypothesis 2 of an age-related decrease of the RPE, there was a significant interaction of task by age ($F_{1,84} = 15.26$, $p < 0.001$, partial $\eta^2 = 0.15$). Separate paired T-Tests (averaged across stimulus novelty, memory and day) showed that both young and older participants benefit from RP (young: $T_{44} = 10.94$, $p < 0.001$, older: $T_{40} = 6.19$, $p < 0.001$) but

that the RPE was larger for young participants than older ones (independent samples $T_{84} = 4.7$, $p < 0.001$), see **Figure 4.3**.

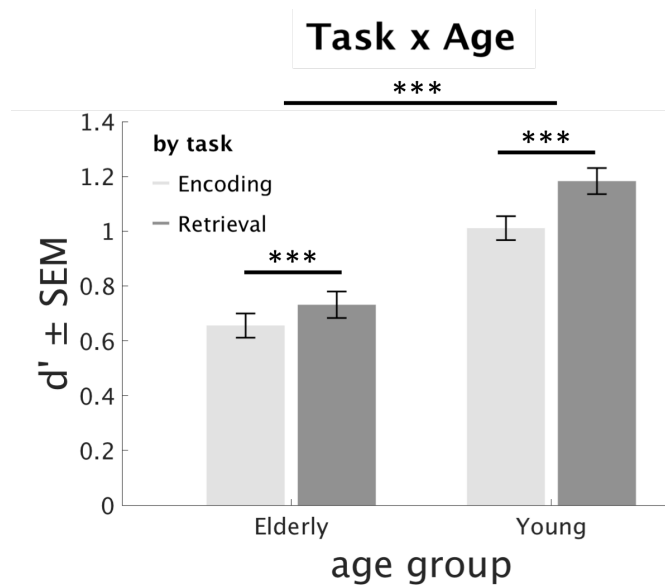


Figure 4.3: Differential effect of retrieval practice on age group. Both young and older participants show the RPE, but it is significantly larger in the Young. *** $p < 0.001$.

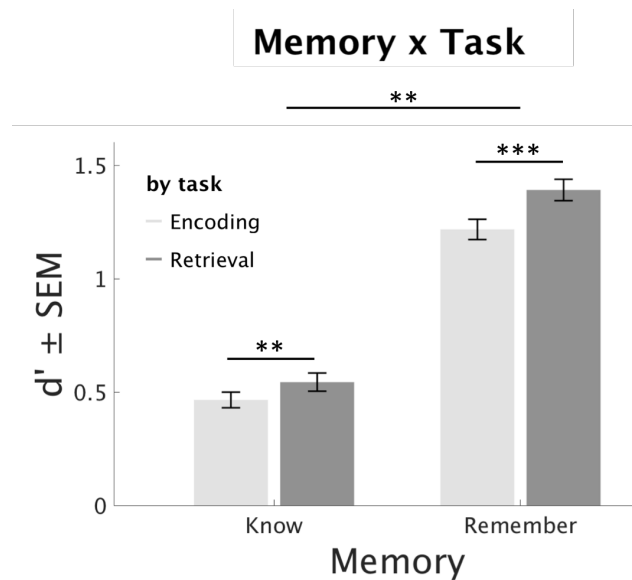


Figure 4.4: Effects of retrieval practice on type of memory. Retrieval increases both types of recognition memory, but this effect is more pronounced for Remember rates. *** $p < 0.001$, ** $p < 0.01$.

Regarding our hypothesis 3, addressing differential influences of recollection and familiarity within the RPE, there was a significant interaction of memory by task ($F_{1,84} = 8.21$, $p = 0.005$, partial $\eta^2 = 0.09$). Dependent samples T-Tests (averaged across stimulus novelty, age and day) showed higher remember and know responses in the RET as compared to the ENC task (know: $T_{85} = 3.57$, $p < 0.01$, remember: $T_{85} = 8.54$, $p < 0.001$). Importantly, the influence of RP was more pronounced for remember responses ($T_{85} = 2.88$, $p = 0.005$), see **Figure 4.4**.

There was a significant interaction of memory by retrieval day ($F_{1,84} = 43.65$, $p < 0.001$, partial $\eta^2 = 0.34$). Dependent samples T-Tests (averaged across stimulus novelty, age and task) revealed that remember responses were higher on Day 1 as compared to Day 7 (know: $T_{85} = 2.29$, $p = 0.024$, n.s., remember: $T_{85} = 13.62$, $p < 0.001$). The difference from Day 1 to Day 7 was stronger for remember than know responses ($T_{85} = -6.658$, $p < 0.001$).

There was an interaction of day by stimulus novelty ($F_{1,84} = 17.28$, $p < 0.001$, partial $\eta^2 = 0.17$). Dependent samples T-Tests (averaged across memory, age and task) revealed that while both old and new stimuli were remembered less on Day 7 as compared to Day 1 (old: $T_{85} = 20.07$, $p < 0.001$, and new: $T_{85} = 18.92$, $p < 0.001$), there was a trend for this effect to be slightly more pronounced for new as compared to old stimuli ($T_{85} = 1.73$, $p = 0.088$).

Finally, there was a significant interaction of memory by stimulus novelty ($F_{1,84} = 128.27$, $p < 0.001$, partial $\eta^2 = 0.6$). Dependent samples T-Tests (averaged across day, age and task) showed no significant difference in know responses for old versus new stimuli ($p > 0.9$). However, remember responses for old stimuli were significantly higher as compared to new stimuli ($T_{85} = 11.41$, $p < 0.001$). The corrected α -level for the post hoc T-tests following significant two-way interactions was $\alpha = 0.0167$.

There were no other significant two-way interactions, including Task by stimulus novelty ($p > 0.15$) or task by day ($p > 0.7$, hypothesis 4). To further investigate the absent effect of task by Day, we conducted a Bayesian paired-samples T-Test to ascertain whether the retrieval benefit was different on Day 1 versus Day 7. Compatible with the ANOVA results, there was only anecdotal evidence for a difference between the two timepoints ($BF_{10} = 2.38$). The retrieval benefit was insignificantly smaller on Day 7 (Day 1: $d'_{\text{Diff}} = 0.3$; Day 7: $d'_{\text{Diff}} = 0.2$).

4.3.3.2.2 Three-Way Interactions

Related to our hypothesis 5, namely that the RPE would be present for old and new items, there was a trend-level interaction (corrected α -Level = 0.0016) between task, stimulus novelty, and age ($F_{1,84} = 7.21$, $p = 0.009$, partial $\eta^2 = 0.08$), see **Figure 4.5**. Post hoc T-Tests (averaged across memory and day) revealed significant RPEs for old and new stimuli for both young and older participants (see **Figure 4.5**). The interaction was driven by a larger difference between RPEs in new, but not old, stimuli in the young adults as compared to older adults. Note, however, that the direct comparison of RPEs for Old and New stimuli within each age group did not reach significance ($p > 0.2$). The Bonferroni corrected alpha-level for post hoc T-Tests following significant interactions was $\alpha = 0.0056$.

There was a significant interaction between day, stimulus novelty, and age ($F_{1,84} = 10.81$, $p = 0.001$, partial $\eta^2 = 0.11$), and a significant interaction of memory, day, and stimulus novelty, $F_{1,84} = 35.82$, $p < 0.001$, partial $\eta^2 = 0.3$. Since these two interactions were not part of our hypotheses and since they did not include Task, they were not followed up by post hoc T-tests. No other three-, four-, or five-way interactions were found.

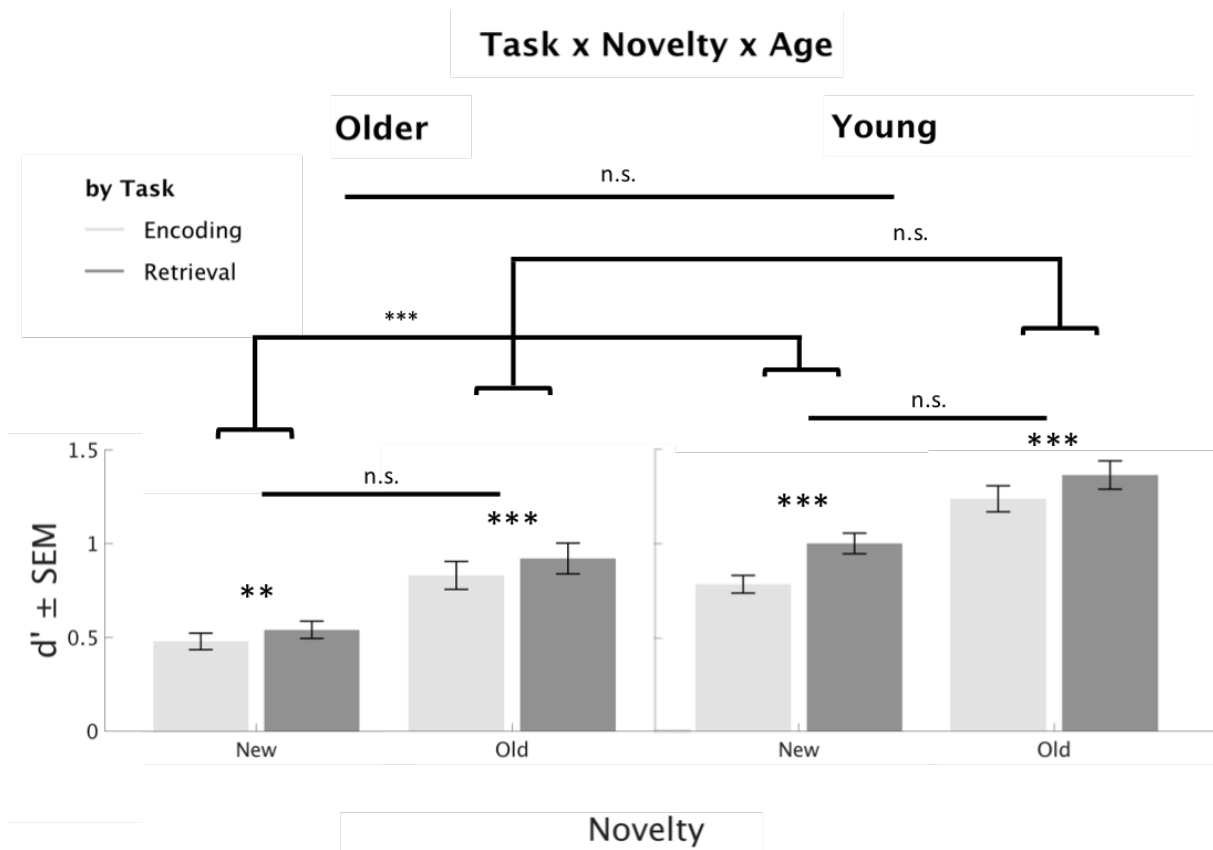


Figure 4.5: Interaction of Task, Novelty, and Age group. The retrieval practice effect was significant for old and new stimuli, and there was no significant difference in strength between old and new stimuli. However, the difference between the RP for old and new stimuli seemed more pronounced in young adults in comparison to older participants. ** = $p < 0.007$; *** = $p < 0.001$.

4.3.4 TCI-Results

Addressing hypothesis 6 on a link between trait Novelty Seeking and the RPE, we calculated Spearman correlations between RP benefit across days, as well as for each day separately, and the TCI “Novelty Seeking” scale – including its subscales (Exploratory Excitability, Impulsiveness, Extravagance, Disorderliness) – in a large sample comprising data from this and a previously reported study (Guran et al., 2019, see methods). There were no significant correlations ($N = 141$, Bonferroni corrected alpha-level: $\alpha = 0.003$). However, a trend for a correlation between the RPEs, averaged across both days ($r = 0.218$, $p < 0.009$), on Day 1 ($r =$

0.181, $p = 0.032$), and on Day 7 ($r = 0.165$, $p = 0.05$), with the TCI’s “Novelty Seeking” subscale “Exploratory Excitability” was found (see **Table 4.3**). Furthermore, there was a trend for a correlation between the RP benefit overall and the subscale “Impulsivity” ($r = -0.166$, $p = 0.049$).

4.3.5 Reduced models and exploratory analyses

In order to avoid the possibility of false positives as a result of overfitting, we conducted a simpler $2 \times 2 \times 2$ ANOVA with the factors memory, task, and age, excluding the factors day and stimulus novelty. It also revealed significant main effects of memory, task, and age (all p -values < 0.001), as well as interactions of task by age, and memory by task (both p -values < 0.01). No other interactions were significant.

Furthermore, since we investigated a broad range of older participants (50–82 years), we separated this group in young-old (50–65, $n = 15$) and old-old participants (60–82, $n = 26$). However, a $2 \times 2 \times 2$ ANOVA with memory, task, and old-age subgroups as a factor did not reveal a main effect of Age ($p = 0.1$). A Bayesian two-sample T-Test revealed that there was moderate evidence for the groups to be the same, $BF_{01} = 3.1$.

Table 4.3: Spearman correlation between TCI Novelty Seeking and RPEs, $N = 142$. * = $p < 0.05$. No significant effects after correction for multiple comparisons ($\alpha = 0.003$).

		TCI Novelty Seeking Subscales				Novelty Seeking ^a
		Exploratory Excitability	Impul- siveness	Extra- vagance	Disorder- liness	
Overall	<i>r</i>	.213*	-.164	.008	.131	.039
	<i>p</i>	.011	.051	.927	.120	.648
RPE s	Day 1 <i>r</i>	.175*	-.116	-.031	.163	.053
	<i>p</i>	.038	.171	.716	.053	.529
Day 7	<i>r</i>	.165	-.127	-.003	-.009	-.005
	<i>p</i>	.050	.132	.976	.911	.957

^aincludes all four subscales

4.4 Discussion

Retrieval practice reliably improves long-term memory, but its temporal stability, relation to recollection/familiarity, aging, and Novelty Seeking has not been systematically investigated. In this study, we could show that the RPE remained stable across a 7- day period, in young as well as older adults, while older adults benefit less from RP. The RPE was present for both previously encountered (old) and novel (new) stimuli, which suggests that enhanced difficulty (or effort) may play a role. Importantly, RP improved both types of memory rates, but the effect was more pronounced for recollection as compared to familiarity indicating a differential effect of RP on recognition memory. Lastly, no convincing evidence was found in favor of a relationship between Novelty Seeking and RP benefits.

As hypothesized (hypothesis 1), RP improved long-term memory in both young and older participants (**Figure 4.3**), which is in line with evidence that RP or testing leads to better memory outcomes compared to restudy or elaborative encoding strategies (Guran et al., 2019; Karpicke & Blunt, 2011; Meyer & Logan, 2013; Rowland, 2014). The RPE has been present in many different settings (word lists, scientific texts, and now pictures) and seems to be quite general. However, whether our findings (age and time) generalize to other stimulus material, or other memory modalities, remains to be studied. In any case, RP seems to be an effective way of improving memory even in clinical populations, such as patients with traumatic brain injury (Sumowski et al., 2014) or multiple sclerosis (Sumowski et al., 2010). In contrast, the effectiveness of RP in other disorders that directly affect dopamine levels in the brain, such as Parkinson's disease, remains unclear.

In terms of age-related modulation of the testing effect, there is no clear consensus in the literature whether RP reliably leads to improvement in older adults or not, mainly due to a

sparsity in studies investigating age in a comparative setting. Some studies find similar improvements in young and older adults (Balota et al., 1989; Meyer & Logan, 2013) while others report a reduction (Guran et al., 2019) or decreased robustness (Tse et al., 2010), of the RPE in older adults. One possibility is that differences in stimulus materials (e.g., words vs. pictures), and paradigms (free vs. cued recall) might explain the differing findings, which at a more general level, reduces comparability between studies. Moreover, aging is associated with increased interindividual differences regarding a variety of cognitive abilities (e.g., Hedden & Gabrieli, 2004), which may have also differed between studies. Although our data do not allow a clear explanation of different findings, it adds evidence in favor of the notions that (a) older adults also benefit from the RP, and (b) this benefit may be reduced in strength (hypothesis 2). The reduced benefit for RP for older adults is particularly notable, considering that in Phase 1 older adults performed slightly better, which suggests they were more attentive to the task.

As hypothesized and previously shown, the RPE was present immediately and 7 days later (hypothesis 4), showing that RP-related memory improvement is fast and stable. This conclusion is in line with the “Fast Route to Consolidation” hypothesis, suggesting that RP leads to fast consolidation by enhancing the integration of new information into preexisting neocortical networks (Antony et al., 2017). Specifically, this mechanism supposedly relies on fast modifications of neocortical representations, which have been shown elsewhere (Brodt et al., 2018; Tse et al., 2007). Importantly, the RPE was not only fast, but also stable over 7 days, suggesting that RP leads to a long-term modification of memory contents, which is in line with previous studies (Bouwmeester & Verkoijen, 2011; Butler & Roediger III, 2007). However, other studies (Roediger III & Karpicke, 2006; van den Broek et al., 2014) have not

found RPEs in response accuracy immediately after restudy/retrieval sessions. This discrepancy between the above mentioned literature and our findings here and elsewhere (Guran et al., 2019) could be due to the stimulus material and response mode: in previous studies, word pairs were learned and actively recalled, while in our study, we used scene stimuli that were shown again to assess memory. The differences between active recall and recognition memory and how they affect the RPE should be compared directly to see whether the different results possibly reflect different underlying processes within the RP paradigm. Investigating the exact neural processes involved, for instance with fMRI or EEG, would be an important step to further understand the RPE. Interestingly, there was no significant interaction of day by age or day by task by age, suggesting that the RPE was stable over time in both young and older participants. Thus, although overall memory performance was lower in older adults (which is compatible with previous research, see Hedden & Gabrieli, 2004; Nyberg et al., 2012), their RPE, did not show more deterioration over the span of the retention interval.

As expected, even stimuli that were novel during RP benefited from retrieval and were better recognized on Day 1 and Day 7 (hypothesis 5). This is compatible with previous work (Chan et al., 2006; Cho et al., 2017; Guran et al., 2019; Herweg et al., 2018) and further suggests that encountering stimuli in a retrieval mode leads to stronger memory traces irrespective of the stimulus Novelty. This cannot easily be explained by the Episodic Context Account or the fast-route to consolidation hypothesis. Instead, it might be explained by differences in difficulty and semantic elaboration: accordingly, retrieval is more effortful than encoding (Pyc & Rawson, 2009; Rowland, 2014), which may lead to deeper information processing as seen by slower response times (Phase 2) and activity increases in brain regions involved in long-term

memory, such as the mesolimbic system (Herweg et al., 2018; van den Broek et al., 2013). Interestingly, post hoc analyses of the trend-level three-way interaction between task, stimulus novelty and age groups suggest that the difference in the RPE for new, but not old, stimuli was more pronounced for young as compared to older adults (**Figure 4.5**). This may relate to age-related functional or anatomical changes of the underlying neural processes involved. Specifically, novelty processing relies on the dopaminergic system (Bunzeck et al., 2014; Lisman et al., 2011; Lisman & Grace, 2005), including the SN/VTA and MTL, which degenerates with age (Bäckman et al., 2006; Bunzeck et al., 2007; Düzel, Bunzeck, et al., 2010). Importantly, RP had a stronger effect on recollection than familiarity rates (**Figure 4.4**, hypothesis 3). These findings further support the interpretation that RP leads to more semantic elaboration of practiced information in comparison to non-practiced information (Carpenter, 2009, 2011), which in turn increases recollection rates (Chan & McDermott, 2007). Indeed, previous studies could show a decrease of alpha and beta oscillations during RP (Guran et al., 2019), or stronger RP for less semantically bound word pairs (Carpenter, 2011), which further suggests a role of semantic processes or semantic elaboration in RP. However, in our study, familiarity rates were also significantly enhanced through RP, indicating that semantic elaboration is not the sole contributor to the RPE (Lehman et al., 2014).

The differential effect of RP on recollection and familiarity is compatible with dual process models of recognition memory and functional imaging studies (Diana et al., 2007; Yonelinas et al., 1996, 2010). They suggest different regions of the MTL to be involved in the two different memory experiences: the hippocampus and posterior parahippocampal gyrus are closely associated with recollection, the anterior parahippocampal gyrus, however, with familiarity (Diana et al., 2007). Therefore, the hippocampus appears to be more critical for

recollection but not for familiarity (Yonelinas et al., 2010). Whether the effect of RP on recollection and familiarity is also mediated via these brain regions, needs to be investigated in future studies with lesion approaches or functional brain imaging techniques such as fMRI. Based on a link between the dopaminergic system, Novelty Seeking, and declarative long-term memory (Benjamin et al., 1996; Chowdhury et al., 2012; Ebstein et al., 1996; Krebs et al., 2009; Schott et al., 2006), we expected a relationship between Novelty Seeking and RPEs (hypothesis 6). However, our results offer no convincing evidence that Novelty Seeking is linked to an individual's benefit from RP. There was a marginal trend for the Novelty Seeking Scale "Exploratory Excitability" to be weakly positively linked to RP benefit. Future studies of the RPE should consider adding the TCI, or other tools to measure Novelty Seeking, to their paradigm. Alternatively, data across studies might be pooled to further increase sample size and power for this apparently small effect, and neuroimaging studies could directly investigate the hypothesized link between RP, Novelty Seeking, and dopaminergic neuromodulation.

In addition to the hypothesized effects, we observed significant interactions between memory and stimulus novelty, memory and day, as well as day and stimulus novelty. Regarding the interaction of memory and stimulus novelty, post hoc analyses indicates significant differences between old and new items only for recollection but not familiarity judgments. This suggests that stimuli that had been encountered more often had more opportunities to elicit elaborate associations that could later be recalled. Post hoc analysis of the significant interaction between memory and day revealed that recollection and familiarity scores were higher on Day 1 as compared to Day 7, but the difference between both days was more pronounced for recollection. In other words, higher recollection memory scores on Day 1 had a sharper decrease (i.e., higher forgetting rate) over time. The feeling of familiarity, which is

supposed to be rather immediate and less resource intensive (Basile & Hampton, 2013; Yonelinas, 2002), was lower as compared to recollection (main effect Memory) but remained more stable from Day 1 to Day 7. From a more general perspective, the interaction of memory and day further underlines dual-process models mentioned above. And lastly, the interaction of day and stimulus novelty was driven by a more pronounced forgetting rate over time for new as compared to old stimuli. This effect can also be explained based on the different numbers of encounters. While new stimuli were presented only once, old stimuli were presented four times, which makes new stimuli more prone to forgetting.

Together, retrieval of information drives declarative memory at an immediate and longer retention interval. RP affects both types of recognition memory, yet the RPE is more pronounced for recollection as compared to familiarity. The RPE is pervasive in the sense that it is visible across the life span, yet older adults benefit less from retrieval as compared to younger adults. As such, our data give novel insights into the temporal stability of the RPE, differential effects of RP on recognition memory types, and they provide further evidence for age-related changes and decreases in the RPE.

Chapter 5 – Retrieval Practice promotes Fast Neocortical Representations in Humans

Currently, this chapter is under consideration for publication. When published, the content and name of the paper might differ from this chapter. Lorena Deuker, Martin Göttlich, Nikolai Axmacher, and Nico Bunzeck are co-authors of this paper.

5.1 Introduction

While repeated study is very common to learning new materials, retrieval practice, or testing, is a more beneficial approach to learning and retaining information (Abbott, 1909; Karpicke & Roediger, 2008; Roediger III & Karpicke, 2006; Spitzer, 1939). Behaviorally, the retrieval practice effect (RPE) is superior to simple and elaborative encoding (Karpicke & Blunt, 2011), effective for different study materials (including word lists, scientific tests, or pictures, see Rowland, 2014), might protect against interference (Kliegl & Bäuml, 2016, but see Hupbach, 2015), is long-lasting (Kromann et al., 2010), present in children (Carpenter et al., 2009; Karpicke et al., 2016, but see Goossens et al., 2016; Moreira et al., 2019), healthy older adults (Guran et al., 2019; Guran et al., 2019; Meyer & Logan, 2013), as well as patient populations (Sumowski et al., 2010, 2014). Despite such a wealth of evidence from behavioral studies and the apparent utility in educational contexts, our understanding of the neural mechanisms underlying the RPE is still scarce.

On a neural level, testing versus restudy leads to increased activity in the striatum, precuneus and medial prefrontal cortex (mPFC) (Herweg et al., 2018), the hippocampus, lateral temporal cortices, and connectivity between hippocampus and the mPFC (Wing et al., 2013). This is

compatible with previous studies on the neural correlates of declarative long-term memory and further suggests that these brain regions are of particular relevance with regard to the RPE. Specifically, the medial temporal lobe, including the hippocampus, entorhinal cortex, and parahippocampal cortex (Squire & Zola-Morgan, 1991), but also the mPFC (Bunge et al., 2004; Frey & Petrides, 2000; Xiang & Brown, 2004) and parietal regions (see Wagner et al., 2005) have long been identified as part of a core network underlying long-term memory encoding and retrieval (Rugg & Vilberg, 2013). Importantly, these brain regions deteriorate with older age, especially temporal and frontal cortices (Hedden & Gabrieli, 2004), and accordingly, declarative LTM is often impaired with age (Nyberg et al., 2012). However, older adults do benefit from RP, but this memory benefit might be reduced as compared to younger adults (Guran et al., 2019; Guran et al., 2020), or less universal and robust (Meyer & Logan, 2013; Tse et al., 2010).

In terms of mechanistic explanations of the RPE, many different – not mutually exclusive – frameworks have been proposed over time (see discussion). Here, we focus on a recent one, namely, the so-called Fast-Route to Consolidation (FRC) hypothesis (Antony et al., 2017). It suggests that RP leads to a rapid online consolidation of information, circumventing longer processes including sleep consolidation. Further, retrieval selectively enhances stimulus-specific neocortical networks and downregulates irrelevant connections, thereby leading to faster and more accurate future retrieval attempts. Importantly, the theory claims that retrieved information should quickly become hippocampus-independent and should, therefore, rely more strongly on neocortical representations.

Direct empirical evidence in humans in favor of the FRC hypothesis is limited. For instance, during free recall of word pair associations, hemodynamic activity within a wide network of

brain regions was reduced for those words that were previously restudied as compared to those that were previously retested (Keresztes et al., 2014); similar but regionally more specific effects have been observed within the anterior cingulate cortex (ACC) (Eriksson et al., 2011). Moreover, Ferreira et al. (2019) found increased activity in the mPFC across retrieval attempts, however, a decrease in HC activation was only noticeable after a longer retention interval of two days. In a study not directly aimed at the RPE (Brodt et al., 2018), rapid, yet temporally stable, microstructural changes within the posterior parietal cortex in the context of a study/retrieval paradigm could be demonstrated (Brodt et al., 2016). These findings further suggest that plasticity related changes in the neocortex, associated with the learning of new episodic information, are not necessarily limited to days, as many theories of systems consolidation suggest (Frankland & Bontempi, 2005; McGaugh, 2000); instead, retrieval may lead to fast changes in neocortical representations.

To test the following hypotheses on a neural level, we employed fMRI in combination with state-of-the-art representational similarity analysis (RSA, e.g. Dimsdale-Zucker & Ranganath, 2018; Kriegeskorte et al., 2008). While univariate fMRI analyses allow conclusions regarding increases or decreases in hemodynamic activity, RSA is a specific type of multi-voxel pattern analysis (MVPA) and provides more direct insights into the representation of information, including memories, on the basis of correlations between pairs of distributed activity patterns. Based on previous findings, we expected the RPE to be driven by enhanced recollection in both age groups, but reduced in the older participants. We also expected retested stimuli in a final retrieval test to be a) less dependent on the hippocampus and temporal cortex, and b) to be more dependent on the neocortex. Both hypotheses were investigated in terms of activation, as well as representational similarity changes.

5.2 Materials and Methods

5.2.1 Sample

Our sample comprised a total of 64 younger (20-30 years of age) and older (51-77 years of age) participants. They were recruited through the Online Recruitment System for Economic Experiments (ORSEE, Greiner, 2015). In addition to the inclusion criteria described in chapter 2.1, participants had to be MRI compatible (have no ferromagnetic implants). We had to exclude nine participants from the initially measured sample (three participants misunderstood the task, one had a MoCA score of <22, one showed very poor performance, one had a brain abnormality discovered during scanning, one showed excessive head movements in the scanner). This reduced our sample to 30 younger (Mean age = 24.5 ± 3.1 , 16 female) and 25 older (Mean age = 61.4 ± 6.7 , 16 female) participants.

5.2.2 Experimental Design

In this study, Phases 1 and 2 were performed as described in chapter 2.2 Experimental Paradigm. For Phase 3, participants were brought into the scanner and viewed the stimuli via a mirror reflecting a screen positioned at the back of the scanner. In the scanner, participants performed the surprise recognition task roughly 15 to 20 minutes after the end of the second phase. Participants gave their answers in the form of source memory judgments, with the options “New”, “Unsure”, “Know”, and “Remember”. The experimental paradigm can be seen in **Figure 5.1**.

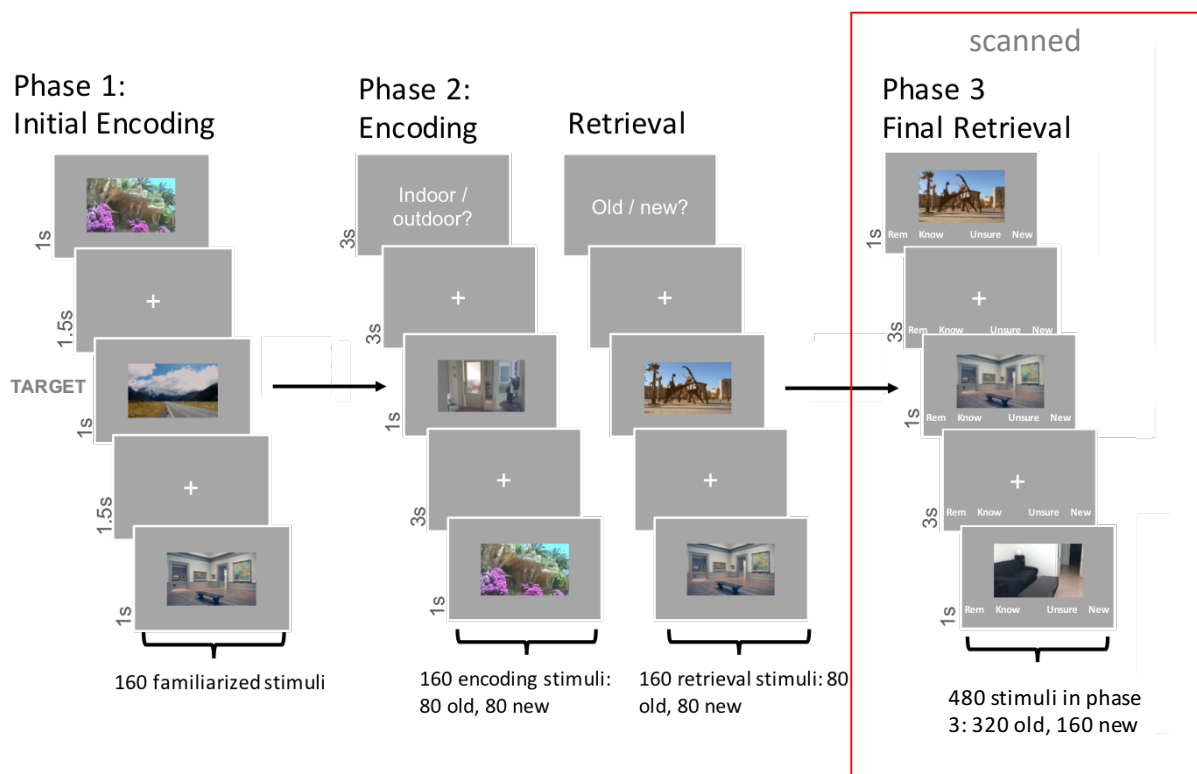


Figure 5.1: Experimental paradigm. In Phase 1, participants were familiarized with 160 indoor and outdoor images by means of a target detection task. These stimuli, intermixed with 160 new stimuli, were presented in Phase 2 in a block design, randomly assigned to either an encoding task (indoor/outdoor categorization), or a retrieval-practice task (old/new categorization). Finally, participants were brought into the scanner and shown all 320 items from Phase 2, intermixed with 160 new stimuli, in a final recall task. Participants had to categorize the stimuli as either remembered, known, unsure, or new. Adapted from Guran, Lehmann-Grube & Bunzeck, 2020.

Phase 3 was conducted in four functional runs, of circa ten minutes each. Participants could break between runs, but also had the chance to break every circa 3 minutes within a run.

5.2.3 Image acquisition and preprocessing

Structural and functional MR imaging was performed at the University of Lübeck, CBBM Core Facility Magnetic Resonance Imaging, using a 3-T Siemens Magnetom Skyra scanner equipped with a 64-channel head-coil. Functional images were acquired applying a single-shot gradient-recalled echo-planar imaging (GRE-EPI) sequence sensitive to blood oxygen level dependent

(BOLD) contrast (TR = 1650 ms; TE = 25 ms; flip angle = 75°; voxel resolution 2.5x2.5x2.5 mm³; 80 x 80 matrix; 60 transversal slices; GRAPPA factor 2 and simultaneous multi-slice factor 2). We recorded four runs with 345 volumes each (plus 6 dummy scans). Structural images of the whole brain using a 3D T1-weighted MP-RAGE sequence were acquired (TR=1900 ms; TE=2.44 ms; TI=900 ms; flip angle 9°; 1x1x1 mm³ resolution; 192x256x256 mm³ field of view; acquisition time 4.5 minutes). Additionally, we acquired resting state runs of each participant (TR = 1650 ms; TE = 25 ms; flip angle = 80°; voxel resolution 2.5x2.5x2.5 mm³; 80 x 80 matrix; 60 transversal slices; GRAPPA factor 2 and simultaneous multi-slice factor 2), however, those data will not be presented in this article. Data were preprocessed using SPM 12 (<http://www.fil.ion.ucl.ac.uk/spm/software/spm12/>, RRID: SCR_007037) and in-house scripts. EPIs were slice time corrected, and realigned to correct for head motion. T1 images were coregistered onto mean EPIs. All images were normalized to MNI space based on normalization parameters derived from a segmentation of T1-weighted images into white matter, gray matter, and CSF using default tissue probability maps. Smoothing was performed with a 6mm full width at half maximum (FWHM) Gaussian kernel. As mentioned above, one participant was excluded due to excessive head motion (exceeding 4mm in each run). For the RSA, preprocessing followed a similar pipeline but neither normalization nor smoothing was performed on the functional data. In addition, T1 images (mapped onto mean EPIs) were segmented using Freesurfer (Fischl, 2012).

5.2.4 Analysis and preprocessing of functional MRI data

MRI data were analyzed using SPM12 (<http://www.fil.ion.ucl.ac.uk/spm/software/spm12/>, RRID:SCR_007037), and custom Matlab scripts, in case of the RSA. For the univariate analysis in SPM, first and second level analyses were performed. On the first level, we included

regressors for the levels of *Task* (study vs retrieval) and *Novelty* (old vs new, i.e. images that were repeatedly presented or only once during the retrieval practice task in Phase 2), resulting in the conditions study-old, study-new, retrieval-old and retrieval-new. Correct rejections, false alarms, guess responses and errors were also modelled on the first level, but not included in the analysis on the second level. The factor memory type (remember vs. know responses) could not be included due to insufficient trial numbers in each condition (see results). Therefore, we averaged the hemodynamic responses across remember and know trials. We used the “FAST” setting for temporal autocorrelation modeling (prewhitening) (Bollmann et al., 2018; Olszowy et al., 2019). Data were high-pass filtered (128s) and motion parameters were included in the model. The four functional runs were concatenated using the inbuilt `spm_concatenate` function, and the data were fitted to the canonical hemodynamic response function (HRF).

On the second level, contrasts of interest were entered into a flexible factorial design with the within-subjects factors task (study vs retrieval) and novelty (old vs new), and the between-subjects factor age group (young vs older) (see Gläscher & Gitelman, 2008). The uncorrected cluster-forming threshold for all analyses was $p < 0.001$, with a minimum of 50 voxels. Clusters of at least 50 voxels, below a familywise-error corrected p-value of 0.05 were considered significant (whole brain, disregarding activations in the cerebellum as we had no particular hypothesis for this part of the brain).

For the multivariate RSA, we calculated weighted 3d images for each trial, based on stimulus onset (weighted mean between volumes, with a temporal lag of 5 seconds to account for the hemodynamic response peak). Systematic effects of motion were regressed out (using the `glmfit` function in MATLAB). The pre-identified regions of interest (ROIs, see below) were

defined with FreeSurfer segmented maps and data from each ROI were extracted from each trial's image. For every participant, voxel patterns of every trial were correlated with voxel patterns of every other trial across all voxels of a given ROI. Then, the correlations were averaged across trials. Data were subsequently analyzed across participants in a linear mixed multilevel model in RStudio, Version 1.1.463 (R Core Team, 2013), which allows to account for additional covariates. Following a stepwise inclusion procedure, we first added random effects (RE) and compared the model to the null model, and then continued adding fixed effects (FE), comparing each new model to the previously best one, using likelihood-ratio tests. FEs and were added in order of their conceptual importance regarding our hypotheses, interactions being added after adding relevant main effects.

5.2.5 ROIs

As described in the introduction, we hypothesized that through retrieval of information, memory traces should become less dependent on the HC and more dependent on the neocortex. Therefore, we performed the RSA on the hippocampus, parahippocampal cortex (PHC) and entorhinal cortex (ERC), frontal cortex, parietal cortex and temporal cortex. The parietal cortex ROI included the supramarginal gyrus and precuneus (Lee et al., 2019, lateral and medial: Jonker et al., 2018, Precuneus: Brodt et al., 2018); the frontal cortex ROI included the entirety of the frontal lobe except for the precentral gyrus; and the temporal cortex ROI included the temporal pole and inferior temporal gyrus, see **Figure 5.2**.

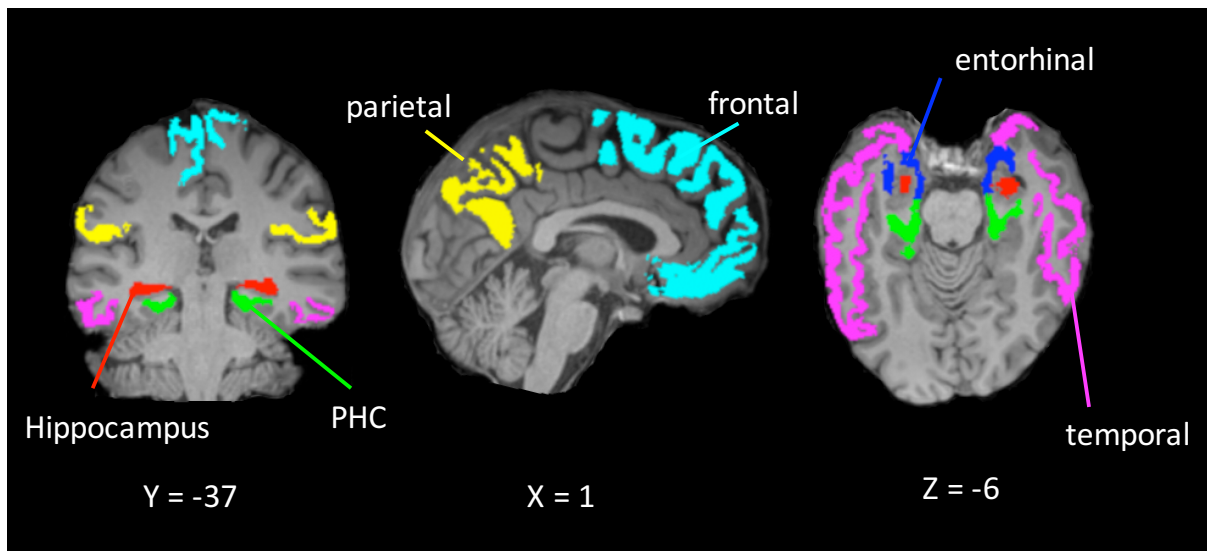


Figure 5.2: Location of the ROIs for the RSA.

5.3 Results

5.3.1 Behavioral results

5.3.1.1 Phase 1

In Phase 1 we compared accuracy (d') and reaction times (RTs) in the target detection task between young and older participants in a two-sample T-Test (Homogeneity of Variance was present in both cases, $p > 0.1$). On average, accuracy was high in both age groups, and there was no significant difference in performance between young and older adults ($p > 0.4$), $d'_{\text{young}} = 5.43 \pm 0.3$; $d'_{\text{older}} = 5.36 \pm 0.38$ (mean $d' \pm$ standard deviation). Similarly, there was no significant difference in RT between both groups, $T_{53} = -1.96$, $p = 0.055$, $RT_{\text{young}} = 556 \pm 71$ msec, $RT_{\text{older}} = 594 \pm 71$ msec (mean RT \pm standard deviation).

5.3.1.2 Phase 2

For Phase 2, we analyzed d' in a 2×2 repeated measures ANOVA with task (study vs retrieval) as within-subjects factor and age (young vs old participants) as between-subjects factor. Note that the factor novelty cannot be analyzed here (see methods). For RTs, we conducted a $2 \times 2 \times 2$ ANOVA including the within-subjects factors task and novelty (old vs new stimuli), and the between-subjects factor age. We tested assumptions of normality of residuals with Kolmogorov-Smirnov Tests and homogeneity of variance with the Levene-Test. As some assumptions (homogeneity of variance, normality of residuals) were violated for some variables, we conducted Bayesian repeated measures ANOVAs alongside the traditional frequentist analysis. In terms of memory accuracy (see **Table 5.1**), the frequentist 2×2 ANOVA revealed a main effect of task ($F_{1, 53} = 270.89$, $p < 0.001$, $\eta^2 = 0.84$), but no main effect of age and no interaction of age by task ($p > 0.5$). A post-hoc T-Test revealed that accuracy was lower in the retrieval task as compared to the encoding task ($T_{54} = 16.55$, $p < 0.001$). The Bayesian

ANOVA confirms this result, suggesting that the best model includes the factor task ($BF_{10} = 4.84e+28$).

Table 5.1: Memory accuracy in Phase 2.

Age group	Task	Mean	SD	N
Young	Study	3.803	0.702	30
	Retrieval	1.795	0.829	
Older	Study	3.889	0.527	25
	Retrieval	1.712	0.701	

Table 5.2: Reaction times in Phase 2.

Age group	Task	Novelty	Mean	SD	N
Young	Study	Old	0.762	0.14	30
		New	0.769	0.15	
	Retrieval	Old	0.886	0.19	
		New	0.917	0.195	
Older	Study	Old	0.819	0.119	25
		New	0.825	0.121	
	Retrieval	Old	0.959	0.147	
		New	0.988	0.134	

For reaction times (see **Table 5.2**), there were main effects of task ($F_{1, 53} = 144.68$, $p < 0.001$, $\eta^2 = 0.73$) and novelty ($F_{1, 53} = 16.86$, $p < 0.001$, $\eta^2 = 0.24$), and an interaction of task by novelty ($F_{1, 53} = 5.81$, $p = 0.019$, $\eta^2 = 0.1$) in the frequentist ANOVA. All other effects were not significant ($p > 0.1$). Post-hoc T-Tests revealed that participants were faster in the study (compared with retrieval) task ($T_{54} = -12.07$, $p < 0.001$), and faster for old (compared with novel) stimuli ($T_{54} = -4.17$, $p < 0.001$). In the Bayesian ANOVA, the best model included both factors task and novelty ($BF_{10} = 2.74e+37$), but not the interaction of task by novelty (task by novelty, $BF_{inclusion} = 0.58$). On a descriptive level, participants were slower in the retrieval task, as well as for new items.

5.3.1.3 Phase 3

In Phase 3, we analyzed d' in a $2 \times 2 \times 2 \times 2$ ANOVA, with memory (remember vs know), task (study vs retrieval), novelty (old vs new in Phase 2) and age (young vs older participants) as factors. Homogeneity of variance (Levene-Test) was present in all but one variable (retrieved new items that were rated as “known”), and normality of residuals (Kolmogorov-Smirnov Tests) was present in all but three variables (retrieved old items that were rated as known in older participants, encoded old items that were rated as known by older participants, and retrieved old items that were rated as remembered by young participants). Because of the large sample size of $n = 55$ participants in total, we still conducted frequentist ANOVAs, but we also conducted a Bayesian repeated measures ANOVA. The best model ($BF_{10} = 3.3e+100$) mirrored the main effects described below, as well as the interaction of memory by novelty. A model that also included the interaction of memory by task also performed very well ($BF_{10} = 2.55e+100$).

5.3.1.3.1 Main effects

The frequentist ANOVA revealed main effects of memory ($F_{1, 53} = 161.83$, $p < 0.001$, $\eta^2 = 0.75$), task ($F_{1, 53} = 34.37$, $p < 0.001$, $\eta^2 = 0.39$), novelty ($F_{1, 53} = 131.4$, $p < 0.001$, $\eta^2 = 0.71$), and age ($F_{1, 53} = 6.59$, $p = 0.013$, $\eta^2 = 0.11$). Furthermore, post-hoc T-Tests revealed higher recollection than familiarity rates ($T_{54} = 12.9$, $p < 0.001$), higher memory scores for old items ($T_{54} = 12.05$, $p < 0.001$), and, importantly, higher memory scores for items previously encountered in the retrieval task (retrieved items; $T_{54} = -7.25$, $p < 0.001$, see **Figure 5.3**). Older participants showed a trend for lower memory scores than younger ones ($T_{53} = 1.49$, $p = 0.071$); thus, despite the significant main effect of age in the ANOVA, the effect in the post hoc T-Test was only significant at a trend level (one-sided two-sample T-Test, see **Figure 5.3**).

D' Main Effects by age group

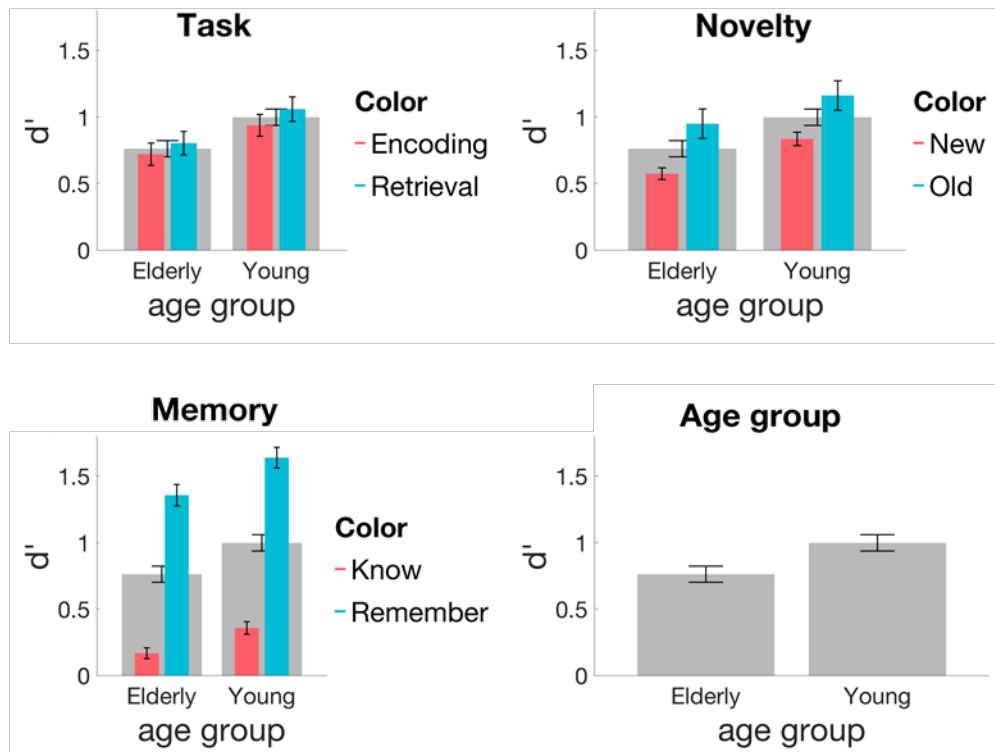


Figure 5.3: Main effects of task, novelty, memory and age group on memory accuracy in Phase 3. Retrieval practice led to higher memory accuracy (A), and old stimuli were recognized more accurately (B). Overall, stimuli elicited more recollection than familiarity (C), and younger participants outperformed older ones.

5.3.1.3.2 Interactions

There was a significant interaction of memory by task ($F_{1, 53} = 17.58$, $p < 0.001$, $\eta^2 = 0.25$), which was driven by a significant difference between study and retrieval for remembered items, but not for known items (see **Figure 5.4a**; $T_{54} = 7.0$, $p < 0.001$, and $T_{54} = 0.72$, $p > 0.4$, respectively). There was a disordinal interaction between memory and novelty ($F_{1, 53} = 106.23$, $p < 0.001$, $\eta^2 = 0.67$); old items were associated with more recollection and fewer familiarity

responses (see **Figure 5.4b**; $T_{53} = -10.0$, $p < 0.001$, and $T_{48} = -4.27$, $p < 0.001$, respectively).

There were no other significant interactions, including task by age (all $p > 0.05$).

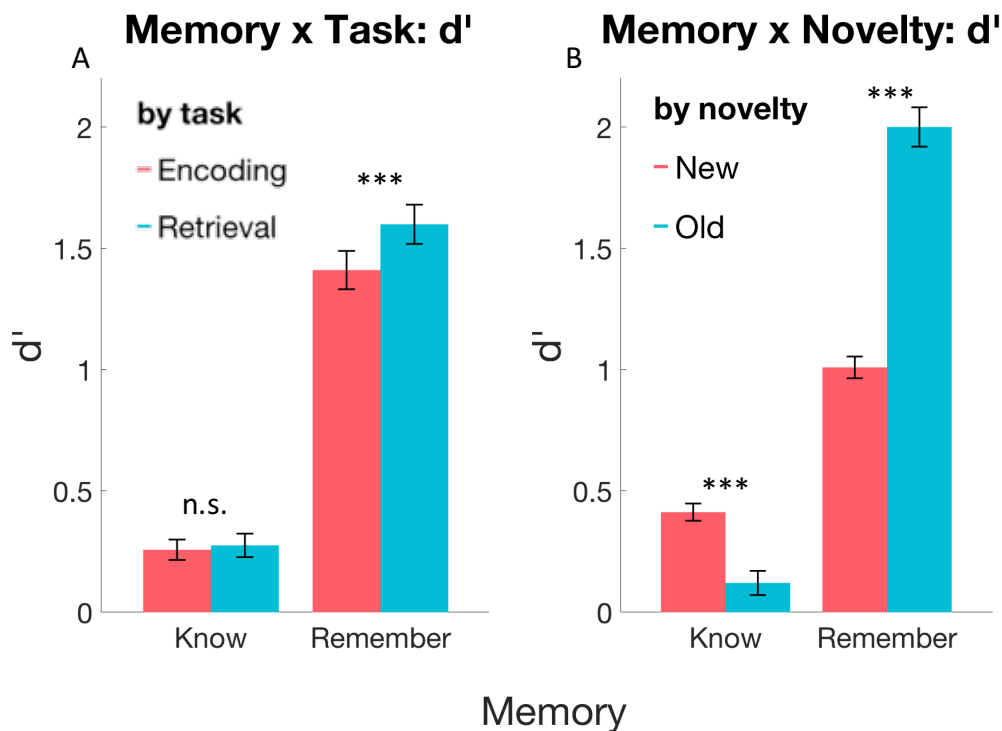


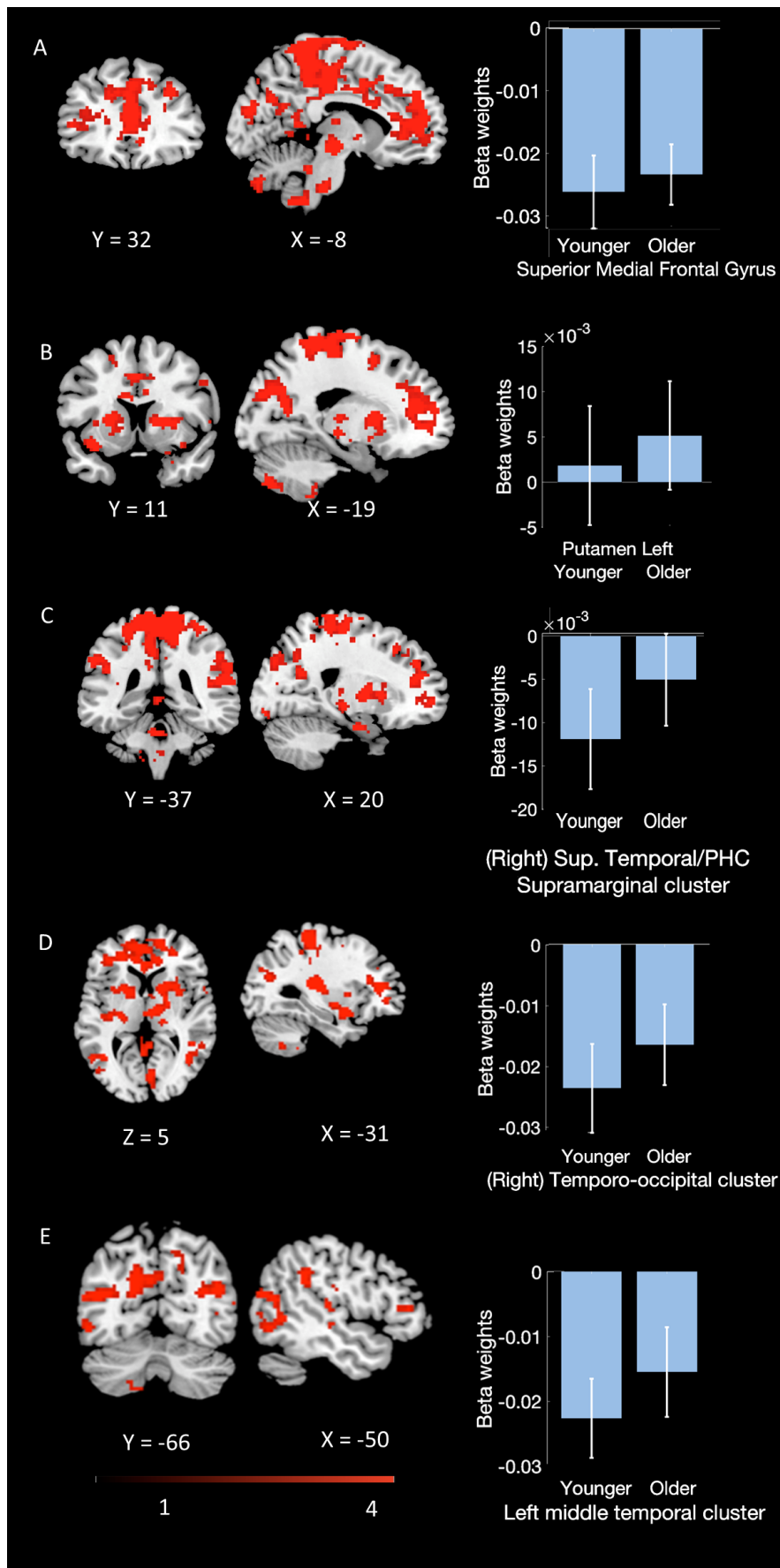
Figure 5.4: A) Interaction of memory type by task. Retrieval practice significantly increased remember responses, i.e. recollection rates, but it had no effect on familiarity. *** $p < 0.001$. B) Interaction of memory by novelty. Remember responses were significantly more associated with old items as compared to new items, while know responses were significantly more associated with new items as compared to old items. *** $p < 0.001$.

5.3.2 fMRI results

5.3.2.1 Flexible Factorial design – univariate analysis

The $2 \times 2 \times 2$ ANOVA with the within-subjects factors task (study vs retrieval), and novelty (old vs new in Phase 2), and the between-subjects factor age (young vs old) revealed main effects of age and novelty. The age effect (contrast older > younger) was associated with widely distributed higher activity for older adults in the bilateral precentral/postcentral gyrus, medial and lateral PFC, temporal cortex, basal ganglia, bilateral thalamus, temporal cortex and

supramarginal gyrus (see **Figure 5.5** and **Table 5.3**). For the opposite contrast (young > older), there was higher BOLD activity in the right frontal inferior operculum, left temporal pole and middle temporal gyrus, HC, parahippocampal cortex, fusiform gyrus, and lingual gyrus (see **Figure 5.6** and **Table 5.4**). The main effect of novelty was driven by stronger activation in central superior motor areas (left and right) for new stimuli in comparison to old ones (see **Figure 5.7** and **Table 5.5**). There were no significant clusters for the factor task ($p_{FWE} = 0.082$, RET>ENC, right supramarginal gyrus), or any of the possible interactions.



(from previous page)

Figure 5.5: (previous page) Age effects: higher activation in older participants as compared to young ones (older > young). There were widely distributed activities in pre- and postcentral motor areas (bilaterally), including the superior medial frontal gyrus (A), left putamen (B), right temporo-parietal cortex, including the PHC and supramarginal gyrus (C), and temporo-occipital cortex (D), as well as the left middle temporal gyrus (E). All activation maps are thresholded at $p < 0.05$ (FWE-corrected at cluster-level using a cluster forming threshold at voxel level of $p < 0.001$). For a full overview of activations in this contrast, see also **Table 5.3**. Bar plots illustrate activity in the cluster.

Table 5.3: Clusters and peaks with stronger BOLD activation in older than in younger participants. Peaks that did not reach significance (FWE-corrected) or lie outside the cerebrum are displayed in grey text.

peak location	peak coordinates (MNI)			statistics			
	x	y	z	p (FWE-corr)	cluster # of voxels	peak p (FWE-corr)	T
(cluster also includes vast areas of PFC and pre and post cenral gyri)				< 0.001		< 0.001	
superior medial frontal R	6	-27	72		712	< 0.001	10.06
superior medial frontal L	8	53	28			< 0.001	7.71
	-12	53	15			< 0.001	7.38
white matter	-34	6	-12	< 0.001	396		8.78
Putamen L	-24	10	8			0.013	5.26
Thalamus R	-24	-17	8	< 0.001	1676	< 0.001	8.67
Pallidum R	10	-12	2			< 0.001	7.91
Insula R	20	-2	8			< 0.001	7.07
Rolandic Operculus L	40	-10	-5			< 0.001	6.56
Putamen R	-44	-27	22			< 0.001	6.22
Parahippocampal R	20	13	5			0.001	5.86
Superior temporal R	20	-7	-22			0.002	5.66
Parahippocampal R	48	0	-8			0.003	5.63
Supramarginal L	26	8	25			0.005	5.49
	-57	-37	35			0.006	5.44
Postcentral gyrus R	66	-14	20	< 0.001	3060	< 0.001	7.49
Superior temporal R	58	-40	20			< 0.001	7.14
Superior temporal R	46	-30	18			< 0.001	6.55
Middle temporal R	46	-54	12			< 0.001	6.41
Superior temporal R	53	-34	18			< 0.001	6.4
Superior occipital L	-17	-70	30			< 0.001	6.36
Insula R	30	-22	20			< 0.001	6.22

Supramarginal R	60	-34	30			< 0.001	6.15
Supramarginal R	60	-27	28			< 0.001	6.1
Cuneus L	-7	-87	20			< 0.001	6.01
Cuneus L	-17	-64	22			0.001	5.99
Cuneus R	8	-84	18			0.001	5.91
Angular R	40	-54	35			0.001	5.89
Cuneus R	8	-84	22			0.001	5.89
Angular R	43	-52	25			0.001	5.74
(vermis)							
Middle temporal L	-42	-70	20	< 0.001	456	< 0.001	6.89
Middle occipital L	-37	-77	28			< 0.001	6.17
Middle temporal L	-54	-64	18			0.001	5.8
Middle temporal L	-54	-62	8			0.054	4.88

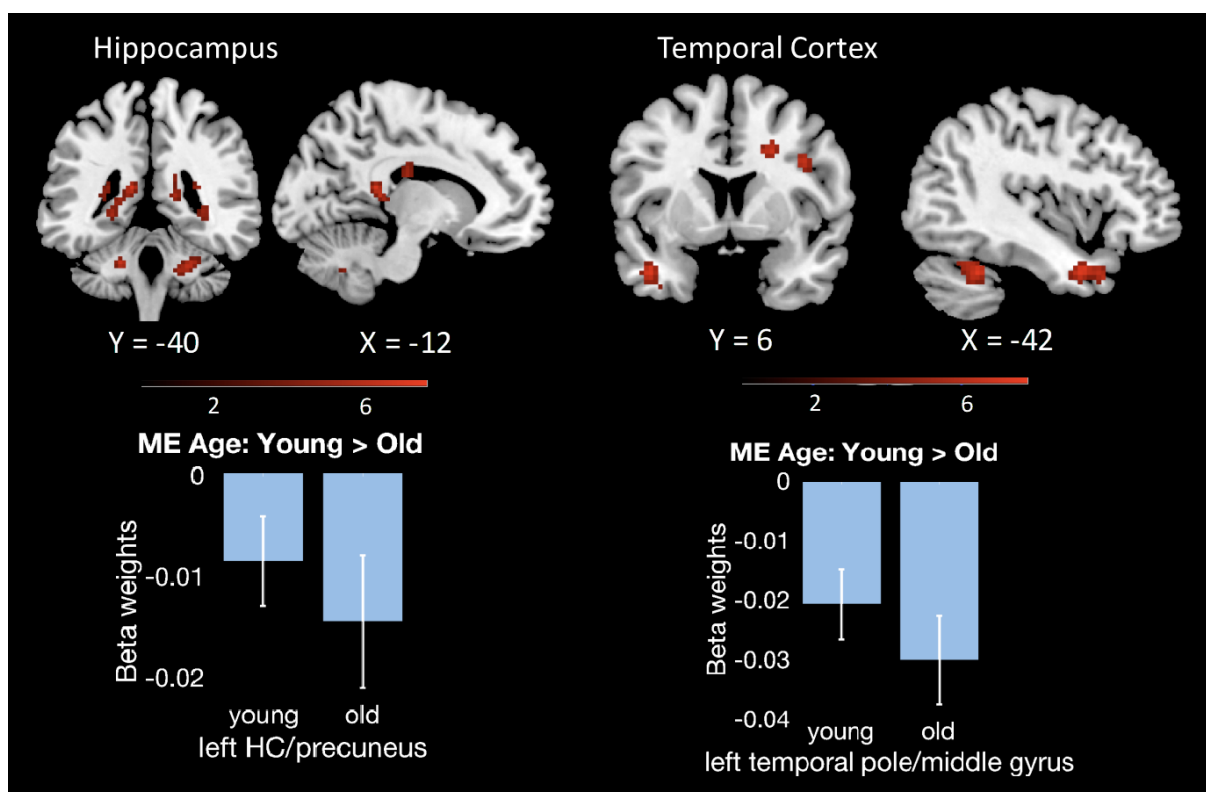


Figure 5.6: Age effects: higher activation in younger participants as compared to older ones (young > older). We found significant activity within the left HC and precuneus (A), and a cluster including the left temporal pole and middle temporal gyrus (B). All activation maps are thresholded at $p < 0.05$ (FWE-corrected at cluster-level using a cluster forming threshold at voxel level of $p < 0.001$). For a full overview of activations in this contrast, see also **Table 5.4**. Bar plots illustrate activity in the cluster.

Table 5.4: Clusters and peaks with stronger BOLD activation in younger as compared to older adults. Peaks that did not reach significance (FWE-corrected) are displayed in grey text.

peak location	peak coordinates (MNI)			cluster		peak	
	x	y	z	p (FWE-corr)	# of voxels	p (FWE-corr)	T
white matter/ventricle R	38	-2	28	< 0.001	496	< 0.001	7.75
posterior cingulum L	-12	-40	15	< 0.001	213	< 0.001	6.71
HC L	-14	-37	20			< 0.001	6.31
Precuneus L	-24	-50	5			0.404	4.26
middle temporal L	-42	6	-30	0.022	96	< 0.001	6.09
temporal pole L	-44	16	-28			0.007	5.41
temporal pole L	-40	18	-32			0.02	5.14
inferior temporal L	-40	-2	-30			0.887	3.8

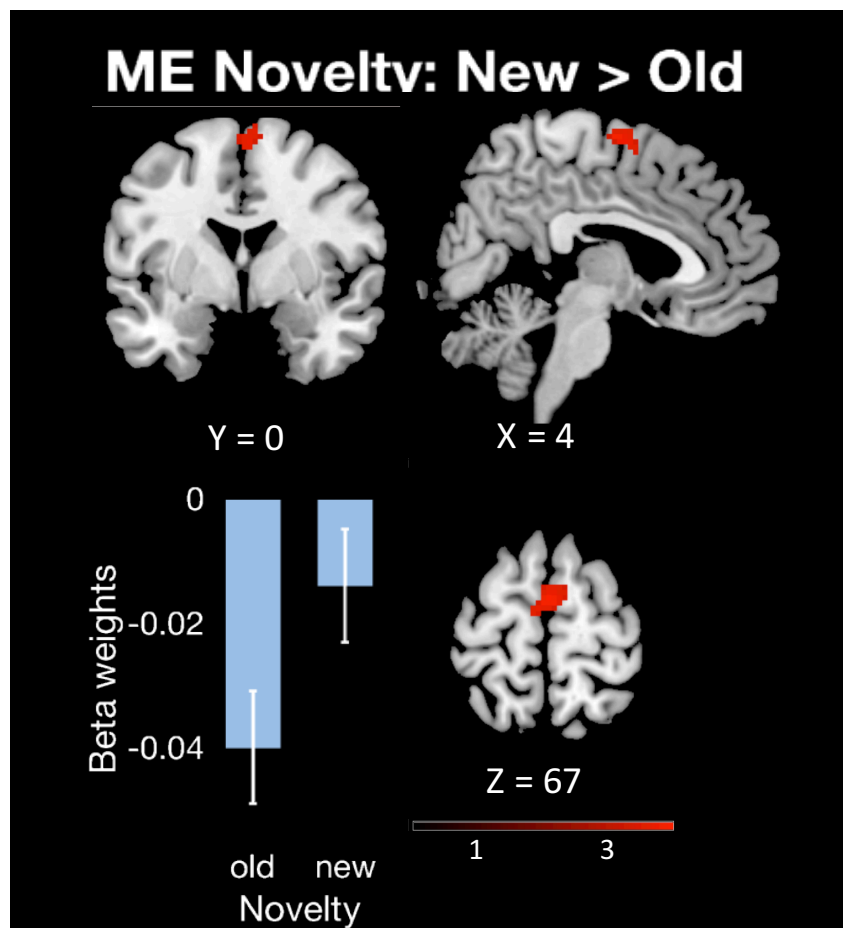


Figure 5.7: Novelty effects. Higher activation for new stimuli, in comparison to old ones, in the bilateral precentral gyrus, see also **Table 5.5**. Beta weights extracted from whole cluster.

Table 5.5: Clusters and peaks with stronger BOLD activation elicited by new in comparison to old stimuli.

cluster location	peak coordinates (MNI)			statistics			
	x	y	z	cluster		peak	
				p (FWE-corr)	# of voxels	p (FWE-corr)	T
supplemental motor area L + R	0	-4	65	0.041	82	0.971	3.99

5.3.2.2 Link between behavior and brain activation – regression analysis

We performed a multiple regression analysis, with retrieval benefit ($d'_{\text{Retrieval}} - d'_{\text{Study}}$ in Phase 3) as a predictor for the difference in brain activation of retrieval versus study (calculated on the first level). In this analysis, we averaged across young and older participants since there was no significant interaction between age and task at the behavioral level and in order to increase the number of participants. We found three significant clusters (see **Table 5.6**), one in superior temporal pole, one in the mPFC (including the medial frontal gyrus), and one in the middle occipital gyrus, see **Figure 5.8**. One outlier with particularly high RP benefit was identified visually and statistically (see **Figure 5.8**, RP benefit value > Mean + 3SD). Correlations in both clusters stayed significant in a Pearson correlation (both $p < 0.025$).

Table 5.6: Areas with higher activity in dependence of an individual's benefit from retrieval practice. The larger the RP benefit, the stronger the activation in each of the areas.

cluster location	peak coordinates (MNI)			statistics			
	x	y	z	p (FWE-corr)	# of voxels	p (FWE-corr)	T
anterior cingulum R+L	6	53	12	0.042	73	0.381	4.63
temporal superior/pole L,							
insula L	-44	-2	-8	0.003	129	0.395	4.61
cuneus R, middle and superior occipital R	28	-82	15	0.039	74	0.853	4.12

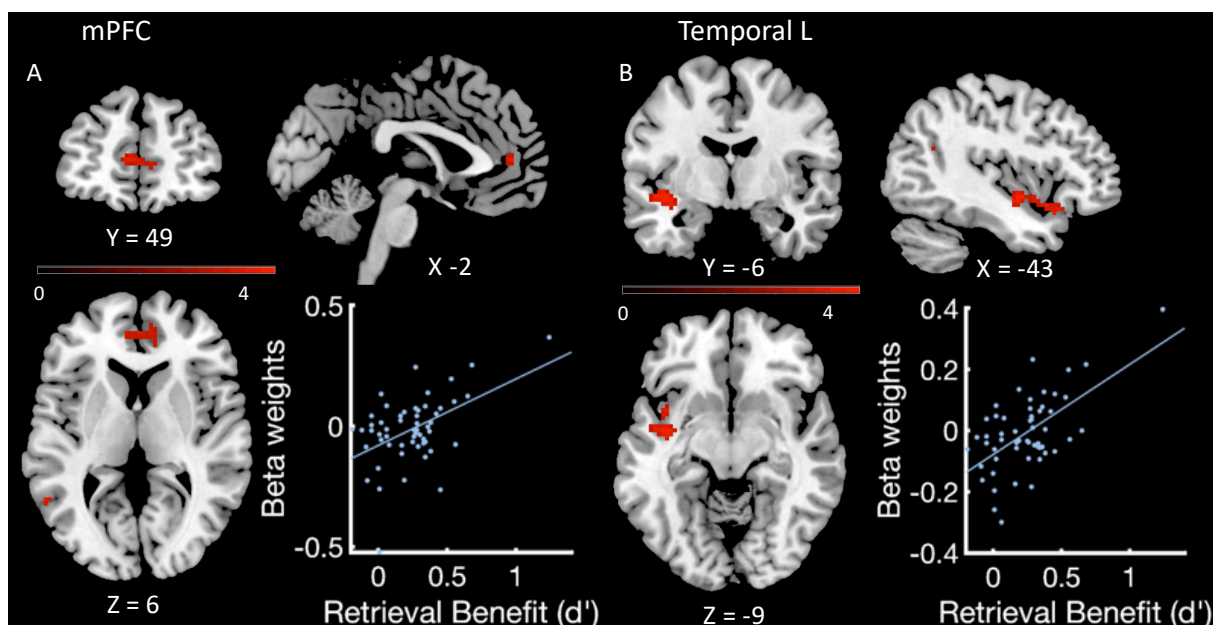


Figure 5.8: Multiple regression between retrieval benefit (x-axis) and activity in mPFC/cingulate (A) and left temporal cortex (B) respectively. RP increased activity in both areas, see also **Table 5.3**. All activation maps are thresholded at $p < 0.05$ (FWE-corrected at cluster-level using a cluster forming threshold at voxel level of $p < 0.001$). Beta weights extracted from cluster.

5.3.2.3 RSA – Multivariate Mixed Model

The models were estimated on unambiguous/even data from correlations within-level of each factor, e.g. study-old with study-old stimuli. The mixed-model analyses revealed participant as a significant random effect (Odd's Ratio (OR) = 38.12, $p < 0.0001$) for pattern similarity (i.e., correlation values extracted for each ROI). In terms of fixed effects, the model including task had a significantly better fit ($X^2 = 14.12$, $p = 0.0002$, $dfs = 4$) for explaining the pattern similarity data. ROI (frontal, parietal, temporal, HC, Entorhinal, Parahippocampal) further improved the model significantly, ($X^2 = 53.52$, $p < 0.0001$). Neither stimulus novelty ($p > 0.5$) nor Age ($p > 0.2$) as fixed effects improved the model significantly in comparison to the more parsimonious random effects model. The model including both task and ROI as fixed effects outperformed

either model containing only one of the factors ($X_{\text{Task}}^2 = 54.13$, $p < 0.0001$, $X_{\text{ROI}}^2 = 14.74$, $p = 0.001$).

We were also interested in interactions between task and ROI, task and age, novelty and ROI, as well as age and ROI, and the three-way interaction between all these factors of interest. The interaction between age and ROI significantly improved the model fit, $X^2=8.02$, $p = 0.005$. Neither the interaction of task by ROI nor of task by age, nor of novelty by ROI improved the fixed effects model significantly ($p > 0.2$, $p > 0.1$, and $p > 0.7$ respectively). Finally, there was a significant three-way interaction of task, ROI and age, which improved the model fit in comparison to the simple fixed effects model ($X^2= 9.45$, $p = 0.002$).

Since the factor novelty had no significant effect in our model, we ran the analysis again, including all possible RSA correlations that were unambiguous for each level of the factor task. Similarly, the model including participants as random effects significantly outperformed the null model ($X^2 = 61.11$, $p < 0.0001$). Task was included as a fixed effect ($X^2= 7.66$, $p = 0.006$, see **Figure 5.9**). Age did not improve the random effects model ($p > 0.4$). ROI significantly improved the model ($X^2 = 70.59$, $p < 0.0001$, see **Figure 5.10**). The interaction of task by age did not improve the model ($p > 0.3$), neither did the interaction between task by ROI ($p > 0.3$). There was a significant interaction of age by ROI ($X^2 = 8.3$, $p < 0.01$). Finally, there was a significant 3-way interaction that outperformed the model with only random and fixed effects ($X^2 = 7$, $p < 0.01$, see **Figure 5.11**). The final model thus included following factors: Task, ROI, and the 3 way interaction of Task by Age by ROI, with participant as a random effect.

Main Effect of Task

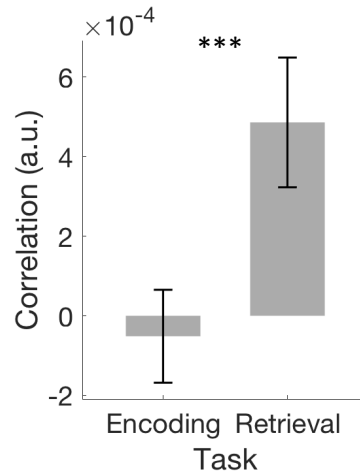


Figure 5.9: Main effect of task in the RSA. Across all ROIs included in the models, stimuli shown in the retrieval task elicited higher similarity at the final recognition memory test than stimuli shown during the encoding task. *** $p < 0.001$.

Main Effect of ROI

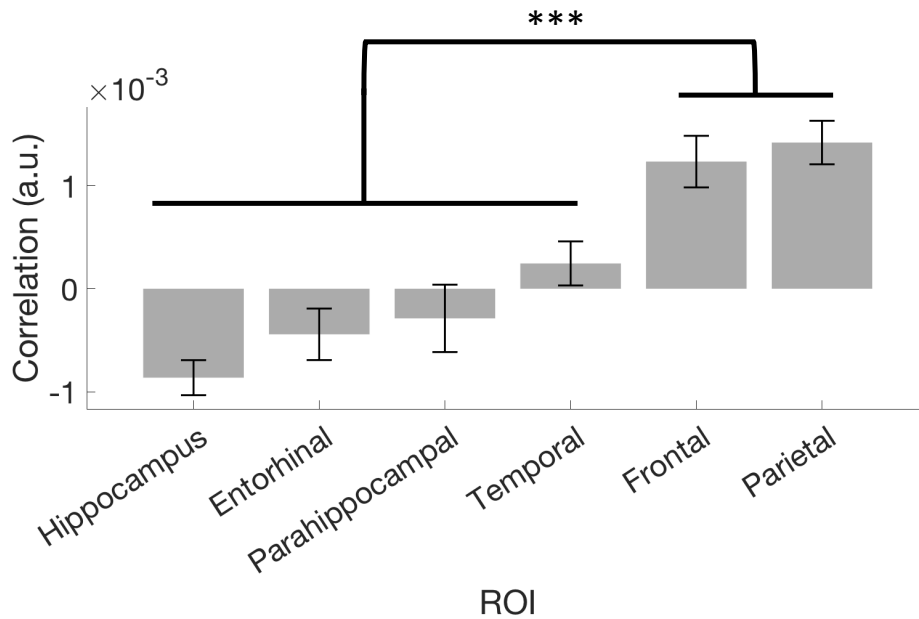


Figure 5.10: Main effect of ROI in the RSA. Frontal and parietal ROIs showed significantly higher similarity than all other regions. Additionally, the hippocampus had significantly lower pattern similarity than the temporal cortex. For further statistical comparisons see **Table 5.7**.

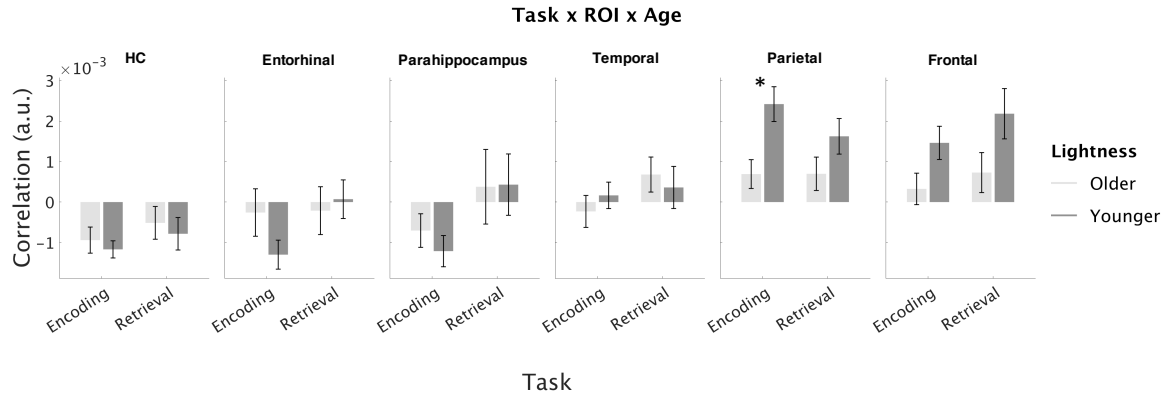


Figure 5.11: 3-way interaction of task, ROI, and age. This interaction was driven by higher pattern similarity in the parietal cortex during the encoding task for young, in comparison to older participants, $p < 0.05$. Overall, older participants showed less similarity modulation with respect to ROI and task than young participants.

Post-hoc T-Tests revealed that stimuli previously encountered in the retrieval condition had a significantly increased similarity in comparison to those from the study condition ($T_{(54)} = -2.15$, $p = 0.018$; test values, see **Figure 5.9**). In terms of ROI, the cortical ROIs of the parietal and frontal cortex had higher similarity than the hippocampus, entorhinal, and PHC (bilaterally), and the temporal cortex had a higher pattern similarity than the HC, but smaller than the parietal and frontal cortices, see **Figure 5.10** and **Table 5.7**.

Because of the number of post-hoc T-Tests (15), the adjusted alpha value is 0.0033 in this contrast. To describe the 3-way interaction, we first visually identified which differences could drive the interaction and then calculated T-Tests on those contrasts, see **Figure 5.11**. On a descriptive level, representational similarity is less affected by task or ROI in older participants. The interaction was driven by lower representational similarity in the parietal cortex during the encoding task in older as compared to younger participants, $T_{(53)} = 2.48$, $p = 0.016$. For this contrast, we only conducted two T-Tests (corrected alpha = 0.025).

Table 5.7: T-Tests for main effect of ROI in the RSA analysis. Alpha-threshold level = 0.003. PHC = Parahippocampus. Significance markers: n.s. exceeds alpha threshold; * $p < 0.003$; *** $p < 0.0001$.

dfs = 54	Hippocampus		Entorhinal		PHC		Temporal		Parietal	
	T	p	T	p	T	p	T	p	T	P
Entorhinal	-1.48	n.s.								
PHC	-2.05	n.s.	-0.41	n.s.						
Temporal	-4.27		-1.92	n.s.	-1.44	0.156				
Parietal	-8.2	***	-5.14	***	-4.51	***	-5.16	***		
Frontal	-7.94		-4.49		-3.82	*	-3.83	*	0.88	n.s.

5.4 Discussion

While the RPE has been studied for over a century, our understanding of its underlying neural mechanisms is incomplete. Using fMRI, we investigated how previously retested and restudied items differ at final recall. Retrieval practice increased recollection-based but not familiarity-based recognition memory, in both young and older adults. In a univariate analysis of the fMRI data, the degree of RP benefit was predicted by activity in the mPFC/anterior cingulate, temporal pole and superior temporal gyrus, irrespective of age group. In line with this observation, a multivariate RSA revealed enhanced pattern similarity following retrieval within memory-related brain regions, including the medial temporal lobe, as well as frontal and parietal cortex. Additionally, pattern similarity was more pronounced within the neocortex (i.e., frontal and parietal lobe) as compared to the medial temporal lobe, including the hippocampus. These findings provide empirical evidence for recent theoretical frameworks by suggesting that the beneficial effects of retrieval practice on recognition memory can – at least in part – be explained by fast changes in neocortical representations. As hypothesized, retrieval of information increased recollection based recognition memory in both younger and older adults (**Figure 5.3** and **Figure 5.4A**). Importantly, this effect could be observed on a very short time scale of about 20 min after retrieval (versus study), which is compatible with our predictions derived from the FRC hypothesis (Antony et al., 2017). It suggests that RP leads to a rapid consolidation of memory (i.e., strengthening of memory content) without the need for slower processes, including sleep-related consolidation (Antony et al., 2017; see below for further explanation). The specific effect of RP on recollection-based memory follows research on the RPE (Guran et al., 2020; Verkoeijen et al., 2011), and provides empirical evidence not only for the FRC hypothesis (Antony et al., 2017) but also for the

“Episodic Context Account” (ECA, Karpicke et al., 2014). According to the ECA, each retrieval attempt is associated with a reinstatement and update of episodic contexts, leading to enhanced recollection-based recognition memory, in line with our finding of increased recollection rates. According to dual-process models (Yonelinas et al., 1996; Yonelinas et al., 2010), recognition can be associated with specific details or associations of the study episode and then lead to recollection, or induce familiarity in the absence of such recollective experience. Further support for dual-process models comes from functional imaging studies, suggesting that the hippocampus and posterior parahippocampal gyrus are closely associated with recollection, whereas the anterior parahippocampal gyrus is more associated with familiarity judgements (Diana et al., 2007). Our fMRI data do not allow us to further address this point, since we had to average across remember and know responses due to insufficient numbers of trials per condition (see below).

Interestingly, the RPE was present irrespective of stimulus novelty in Phase 2. That means, not only familiar but also novel stimuli that were shown in the context of retrieval were remembered better as compared to novel items shown in the encoding context. This is compatible with previous work (Jacoby et al., 2005) including our own (Guran et al., 2019; Guran et al. 2020; Herweg et al. 2018), suggesting that not successful retrieval but retrieval mode per se boosts memory performance (Rowland, 2014). Whether this effect relates to enhance effort (Pyc & Rawson, 2009; Rowland, 2014) or other processes such as fast mapping (Sharon et al., 2011) remains to be investigated.

At the neural level, a univariate analysis revealed differences in activation between old and new stimuli (novelty in Phase 2), as well as for younger versus older adults. Briefly, new stimuli increased activity in the bilateral precentral gyrus, an area typically related to motor functions

(Stippich et al., 2002; Yousry et al., 1995) but not novelty processing. In our study, the responses to novel stimuli were slower than to familiar stimuli (data not shown), thus delaying the motor response, and therefore activity in the motor cortex. Novelty usually elicits activation in parts of the mesolimbic system, especially the SN/VTA (Bunzeck & Düzal, 2006), as well as the hippocampus (Kumaran & Maguire, 2009; Wittmann et al., 2007). However, these regions were not active in our novelty contrast. This might be related to the fact that “novel” stimuli had already been seen once in Phase 2: these stimuli were therefore not truly novel, but less familiar than “old” stimuli, that had been shown three additional times in Phase 1.

In terms of age effects, younger participants had stronger activation in the left HC, Precuneus, middle temporal gyrus, and temporal pole (see **Figure 5.6**). All of these areas have been associated with memory processes (see main text), which is in line with the behavioral finding that younger adults generally showed higher memory accuracy than older adults. In the reverse contrast, older adults exhibited stronger activation in vast parts of the motor cortex, left Putamen and mainly right-hemispheric temporal areas. In aging brains, stronger activation of brain areas has been linked to compensatory processes (Cabeza et al., 2018), and general reductions of specificity, especially in prefrontal regions (Stippich et al., 2002; Yousry et al., 1995). Specifically, to retrieve items in Phase 3, older adults appear to recruit further brain areas, especially on the right hemisphere, while still not reaching the younger adults’ performance level and exhibiting de-differentiation patterns in terms of hemispheric asymmetries (Cabeza, 2002; Reuter-Lorenz & Park, 2010).

Importantly, a multiple regression analysis across all participants showed that the RPE was predicted by activity in the mPFC, anterior cingulate cortex, left superior temporal gyrus and

temporal pole, with stronger activation in these areas relating to larger RP benefits on an individual level. According to traditional models of systems consolidation, the hippocampus initially stores novel information; over time, through post-learning reactivation in resting periods and sleep, this information is being transferred to and stored within the neocortex including the PFC (Deuker et al., 2013; Frankland & Bontempi, 2005; McGaugh, 2000; Zhang et al., 2018). Importantly, this view has recently been challenged by the observation of rapid and temporally stable microstructural changes within the human posterior parietal cortex after learning (Brodt et al., 2018). Moreover, the mPFC is not only involved in sleep-related consolidation processes (Euston et al., 2007), but also in the retrieval of recent and remote memories (Gonzalez et al., 2013). Therefore, neocortical brain regions, including the parietal cortex and mPFC, appear to be an essential part of a rapid learning system. Our data further extend and specify this view by demonstrating a close relationship between neocortical brain regions (including the mPFC) and the positive effects of retrieval practice. The superior temporal gyrus, on the other hand, is essential for visual recognition memory (Nakamura & Kubota, 1995), as tested here, and its anterior parts for semantic processing (Tsapkini et al., 2011; Visser & Lambon Ralph, 2011). While this latter observation may fit to the Elaborative Retrieval Hypothesis (Carpenter, 2009, 2011), according to which RP enhances semantic elaboration and improves memory accuracy, it is, at first glance, at odds with the fact that our RPE was driven by enhanced episodic-like (remember rates) but not semantic-like memory (familiarity rates). However, episodic and semantic information processing are not independent, and semantic elaboration, therefore, may positively influence episodic memory (Staresina et al., 2009).

The multivariate RSA provides further evidence that RP leads to rapid changes in the neural representation of information. Specifically, items previously encountered in a retrieval context were, after circa 20 min, significantly more similar in their neural representations in a priori defined regions of interest (**Figure 5.2**), as compared with items that were encountered in the encoding task. Importantly, this effect was (statistically speaking) independent of ROI, and, therefore, observed across all cortical and subcortical regions, including the frontal, parietal, and temporal lobes, as well as the hippocampus and adjacent areas. This was unexpected, since we hypothesized a dissociation through retrieval practice in prefrontal and parietal regions as compared to the MTL, as suggested by the FRC hypothesis. Therefore, our findings are only partly compatible with the FRC hypothesis, and they suggest that immediate retrieval of retested information is not completely hippocampus-independent, but rather relies on changes in neural representations across the brain. To further investigate this issue, future studies should systematically vary the length of the retention interval between retrieval (Phase 2) and final recall (Phase 3). In any case, rapid increases in pattern similarity in frontal and parietal brain regions might reflect the integration of information into existing cortical networks, and therefore underlie the retrieval practice effect. This is further supported by overall higher pattern similarity (i.e. irrespective of task) within frontal and parietal regions as compared to the MTL (**Figure 5.10**), since preexisting memory networks (or schemas) should be primarily represented in neocortical areas (Tse et al., 2007; Van Kesteren et al., 2012).

Contrary to our hypothesis and own previous work (Guran et al., 2019; Guran et al., 2020), the RPE did not differ between age groups (**Figure 5.3A**). In other words, although older adults had overall reduced memory scores in comparison to their younger counterparts (**Figure 5.3D**), their RPE was unimpaired. One explanation is that this fMRI study had specific inclusion

and exclusion criteria (see methods), which may have led to a specific sample of healthy and high functioning older adults. At the neural level, the absence of age-related impairments in the RPE was paralleled by equally enhanced representational similarity through retrieval in both age groups (i.e. there was no significant interaction between task and age). The only specific and significant age effect in the RSA was driven by a decreased pattern similarity for re-studied trials in the parietal cortex in older as compared to younger adults (**Figure 5.11**). While this does not directly speak to the neural processes underlying the RPE, it provides further evidence for a role of the parietal cortex in study and retrieval processes, which are sensitive to age-related changes (see e.g. Maillet & Rajah, 2014; Uncapher & Wagner, 2009). In summary, the retrieval practice effect is characterized by rapid and specific increases in recollection-based recognition memory and changes in neural activation in both young and older adults. In particular, activity in the mPFC/anterior cingulate and left anterior temporal lobe predicted the increase in recognition memory through retrieval, suggesting an involvement of brain regions previously associated with semantic processing, fast learning, and memory consolidation. Furthermore, RP increased pattern similarity within the frontal cortex, parietal cortex and MTL. As such, our data extend current theoretical views of the RPE and provide empirical evidence that retrieval practice involves fast changes in neocortical representations of information.

Chapter 6 – General Discussion

In the following chapter, I discuss the findings of this dissertation that were presented in chapters 3 to 5, and their meaning for the retrieval practice effect. I will address the general aims described in chapter 1.5. in turn. First, concerning the question of the psychological bases of RP, I will evaluate the behavioral results from the here presented studies. Second, regarding the neural bases of RP, the second aim of this dissertation, I will evaluate the findings from chapters 3 and 5. Third, I will address the final aim by weighing all of the evidence garnered through our studies in light of the different theories of RP. I will come to the conclusion, that all the theories described in chapter 1.4 have their own merit and are supported by parts of the data. Importantly, none of the presented explanatory accounts of RP are mutually exclusive, leading me to propose a comprehensive explanatory model of RP, which identifies the effect RP has on the three main processes of memory: encoding, consolidation and retrieval. Finally, I will turn to the strengths and limitations of this dissertation as well as possible future research avenues.

6.1. The psychological bases of the retrieval practice effect in younger and older adults

In all three studies, we were able to find the RPE in both younger and older adults. RP consistently increased memory accuracy for the pictorial stimuli we used compared with restudy in the encoding task. Regarding the presence of RP in older and younger adults, we can conclude two things: First, retrieval practice is a viable learning technique to improve memory for older adults. In all of our studies, older adults showed improved memory for retested in comparison to restudied items. This suggests that RP at least partly relies on memory processes and brain areas that are less susceptible to age-related change, or that

age-related compensatory mechanisms are able to make up for the loss of function. Second, the degree of RP-related improvement is somewhat dependent on the participants' age. Indeed, in the first and second study, older participants showed reduced RPEs compared with young participants, and the individual amount of RP benefit correlated with alpha-beta power reduction in older adults. Particularly, older adults that showed a more youth-like power pattern also benefitted more from RP. In the third study, however, we did not find a difference in the RPE based on age group; older adults showed similar increases in memory accuracy for retrieved items as did younger adults. Importantly, in this study, older adults also did not perform conclusively worse than younger participants: while the overall F-contrast (ANOVA) revealed a main effect of age, this was not significant in a one-tailed T-Test ($p > 0.05$), suggesting that the sample of older adults in our final study might have, in comparison to the other studies, been comprised of high performing participants. In general, across our studies, average older adults benefit less from RP than young adults. This is particularly important, as it suggests that RP might be employed to help pre-clinical older adults manage their daily memory demands. Furthermore, the amount of benefit from RP might be indicative of general maintenance of select processes underlying RP in older age, such as alpha-beta power decreases, but this remains to be investigated in future studies. The presence of RPEs in older adults also suggests that RP relies on facets of memory that are only partially affected by age-related changes, and RP thereby offers an easy-to-implement way to improve memory – which is crucial for quality of life – in older adulthood.

Regarding our second question, relating to the temporal dynamics of the RPE, we found that the presence and even the strength of the RPE was not influenced by the length of the retrieval interval between Phase 2 and Phases 3/4. This finding is noteworthy, because it means that

even non-verbal materials benefit from retrieval practice, and further, that these RP benefits are very fast, present after a retention interval of approximately 5 minutes, but also temporally stable, being present up to 7 days later. The consistent evidence for fast RP benefits only a few minutes after Phase 2 lends support to the Fast-Route to Consolidation hypothesis, which states that RP leads to quick memory improvement (via consolidation mechanisms, Antony et al., 2017).

Considering our question regarding the relative contributions of RP to recollection and familiarity, chapters 4 and 5 consistently showed that RP improves recollection, rather than familiarity. While RP was also shown to improve familiarity in chapter 4, it did so to a much lesser degree compared with improvements to recollection ratings. Recollection is thought to reflect episodic memory components, thus suggesting that RP particularly enhances episodic memory for information. This finding lends support to the Episodic Context Account (ECA, Karpicke et al., 2014), which predicts RP to increase episodic, and not semantic memory. Importantly, the increase in recollection was present even in older adults, although older participants generally show worse recollection and more familiarity (Duarte et al., 2006; Spencer & Raz, 1995). This furthers the idea that RP increases memory to a degree independently from mechanisms, brain regions, and functions that usually exhibit age-related changes. Heightened RP-induced recollection rates in older age suggest that RP is a learning technique through which episodic memory, which is often impaired in older age, can effectively be enhanced. This finding is very important, because it identifies RP as a viable candidate for interventions designed to help aging individuals to cope with the demands everyday life puts on memory.

It should also be noted that all our findings relating to the RPE were largely independent of stimulus novelty. Both familiarized stimuli (seen in Phase 1) and stimuli that were encountered for the first time during Phase 2 benefitted from being presented in a retrieval context. This raises the question whether the RPE is linked to a sort of “retrieval mode”, which, maybe through heightened task demands, leads to an increase in memory for retested items. In our second study, chapter 4, we found evidence that the RPE for novel stimuli was reduced in older adults (task by novelty by age interaction). The general reduction of the RPE in older adults is thus more strongly driven by a reduction of RP for novel stimuli, which is in line with findings suggesting that brain areas related to novelty processing are impaired in older adults (Bunzeck et al., 2007). However, our findings suggest that the RPE is largely independent of stimulus novelty. This finding is not predicted by the FRC hypothesis (Antony et al., 2017), and cannot be accounted for within the ECA (Karpicke et al., 2014). However, a novelty independent RPE supports the ERH (Carpenter, 2009, 2011), which assumes that retrieval practice generally leads to an increased elaboration of encoded material.

6.2 The neural bases of retrieval practice

6.2.1 RP and electrophysiological time-frequency mechanisms

Based on the involvement of particular frequency bands in human memory, especially theta (Herweg et al., 2020), we investigated the frequency bands which play a role in the RPE, and by doing so, identified possible underlying neural processes of RP. Aside from memory and theta-power, other frequency bands have been linked to specific cognitive functions: alpha-power to attention (Klimesch et al., 1998) and semantic memory processing (Klimesch, 1999), beta to long-range synchronization across distant brain regions (Benchenane et al., 2011),

making these three frequency bands and their related mechanisms potential candidates to explain the RPE.

Importantly, retrieval practice was associated with a decrease in alpha-beta power (circa 9-30 Hz). As alpha power decreases with increased task demands (Fink et al., 2005), the change in neural oscillations could be at least partly attributed to the comparatively much more demanding retrieval task. Indeed, retrieval effort (or desirable difficulty, Bjork & Bjork, 2011) has been proposed as an explanatory mechanism for RP benefits. Our results showing the involvement of alpha-beta power in RP support these explanatory accounts. Furthermore, alpha-beta activity has been linked to semantic processing (Klimesch et al., 1997; Klimesch, 1999). Our results suggest that RP might selectively target semantic elaboration mechanisms to enhance declarative memory, which lends support to the ERH proposed by Carpenter (2009, 2011).

Furthermore, in older participants alpha-beta power for retrieved as compared to re-encoded stimuli was even lower, and this age group simultaneously benefitted less from RP. Crucially, the extent of the decrease in power driven by RP predicted the RP benefit on an inter-individual level in older adults, i.e., those older adults that had more youth-like neural oscillatory signatures (i.e., smaller differences between encoding and retrieval alpha-beta power) simultaneously benefitted more from RP. These findings suggest that alpha-beta power decreases are instrumental in the RPE, and might reflect effort and semantic processing. Indeed different subspectra of the alpha-frequency band have been linked to separate memory functions, with the lower alpha band linked to effort and attention and the upper alpha band (Klimesch et al., 1997) to semantic memory processing. The results from our sample demonstrate that RP is an effortful task (low alpha decrease, and lower accuracy in

the retrieval task in Phase 2, compared with the encoding task). RP also involves semantic memory processing (high alpha power decreases), and the extent of semantic memory processing, operationalized through the alpha-beta decrease, predicts the resulting RP benefit in older adults. The involvement of semantic memory in RP is predicted by the ERH, which views semantic elaboration as the driver behind the RPE.

6.2.2. Brain activation markers and activation patterns in RP

In the final study of this dissertation, we investigated predictions made by the Fast-Route to Consolidation hypothesis. While the simple univariate GLM analysis did not offer insights about RP, in a multiple regression, we found that, irrespective of participant age, RP benefit was predicted by activity in a range of brain areas: the mPFC, anterior cingulate, left superior temporal gyrus, and temporal pole. These brain areas have been linked to memory and consolidation (Euston et al., 2007), fast learning (Gonzalez et al., 2013), and semantic memory processing (Tsapkini et al., 2011; Visser & Lambon Ralph, 2011), lending support both to the FRC hypothesis and the ERH. This finding is particularly important, as it suggests that RP does not just affect episodic, or semantic memory, but likely influences both types of declarative memory. In addition, the results suggest the involvement of consolidation-like processes that directly correlate with the amount of benefit experienced from RP.

More support for both the FRC hypothesis and the ERH came from our RSA findings. RP increased pattern similarity compared with restudy, in all of our ROIs. Pattern similarity expresses the degree of representational overlap, or distinctness, of different items. Stimuli that were encountered in a RP context were more similar to one another, than stimuli seen in the re-encoding context. As is assumed by the FRC hypothesis, pattern similarity increases could reflect the integration of new information into pre-existing cortical networks. This idea

is supported by our finding that pattern similarity was stronger in cortical regions, such as the parietal and frontal cortex: in consolidation accounts in general, and the FRC in particular, memory shifts to cortical areas by consolidation processes (Antony et al., 2017; Barry & Maguire, 2019). This integration of RP stimuli into pre-existing cortical networks might be akin to semantic elaboration, as has been proposed elsewhere (Tse et al., 2007, see above), and is also suggested by our previous findings and the activation in the superior temporal gyrus and temporal pole (Tsapkini et al., 2011; Visser & Lambon Ralph, 2011). However, the increase in pattern similarity between items in the RP context as compared to the encoding task, might also be related to a similarly updated temporal context, as is posited by the ECA.

The lack of an interaction of task by ROI is surprising in the account of FRC. However, a decrease of hippocampus dependence might still take place after longer retention intervals than were employed in the third study and could therefore be investigated in paradigms that increase this retention interval. Here, it would also be important to choose a retention interval that is not too long, in order to avoid confounds from offline sleep-dependent consolidation, as the FRC account clearly specifies that RP should lead to fast online consolidation. Whether such fast decreases in hippocampal activity are likely or even possible to occur at all is an important question that might be answered by employing the RPE, to better understand the underpinnings of long-term memory. Indeed, the concept of hippocampus-independent memories after consolidation is debated (Barry & Maguire, 2019; Frankland et al., 2019).

To summarize, through a combination of univariate regression and multivariate mixed models, we were able to find evidence in favor of the FRC account, but also found links between behavior and brain activity that might reflect semantic processing, which supports the ERH. The findings suggest that RP relies on both episodic and semantic memory, and

involves consolidation mechanisms in line with the FRC account, as well as the ERH, and to a lesser extent, the ECA.

6.3 Implications for explanatory theories of retrieval practice

The third aim of this dissertation was to evaluate our results in light of the existing RP theories.

The three main accounts I will be focusing on are summarized in chapter 1.4.

Firstly, the ERH assumes that RP strengthens memory through enhanced semantic elaboration at retrieval, similar to elaborative *encoding* processes. This means that semantic (memory) processes should play a role in RP: indeed, this is what we found in chapters 3 and 5. Alpha-beta oscillations (chapter 3) have been linked to semantic memory processing (Klimesch et al., 1997; Klimesch, 1999) and activation in the superior temporal gyrus as well as the temporal pole (chapter 5) have also been linked to semantic memory. Furthermore, the ERH is the only theory that can directly explain the often reported finding of a RPE for novel stimuli (Chan et al., 2006; Cho et al., 2017; Herweg et al., 2018, and see the here reported studies), as semantic elaboration of a target word does not depend on that target being familiar. Previous studies focusing on the ERH were unable to find supporting evidence for this theory, and I argue the reason may be that those studies focused on overt elaboration, i.e. an elaboration similar to narrative encoding methods such as inventing stories with the to-be-remembered items. The ERH remains indeed unclear about whether semantic elaboration is a conscious and overt process, or rather an automatic and subconscious one. Our findings suggest that semantic elaboration takes place, and does so in an automatic fashion: participants neither reported that they employed semantic elaboration techniques nor was there, arguably, enough time

for them to narratively elaborate and successfully categorize the images in a timespan of just over two seconds.

Secondly, the FRC hypothesis makes clear predictions as to the process involved in RP, namely *consolidation*, as well as to the brain areas involved in processing items from an RP context. In this theory, RP items should elicit less hippocampal, and more neocortical activity as well as enhanced integration of RP information into neocortical networks. Our behavioral results, which consistently show fast RPEs, present after only a few minutes between RP and final recognition test, support the idea of fast processes. In terms of what these processes might be, our neuroimaging results from chapter 5, suggest that the RPE was linked to increased activity in a network of regions involved in fast memory formation and consolidation (mainly the mPFC/anterior cingulate), giving evidence that fast consolidation processes are playing a role in the recognition of previously retrieved items. The FRC hypothesis directly predicts integration of retrieved information into neocortical networks, and our multivariate approach showed that RP increases neocortical representations by increasing pattern similarity across the brain, speaking in favor of this hypothesis. A small caveat for this theory is the absence of evidence for a decrease of hippocampal involvement for RP memories. One of the core features of consolidation is that this process reduces the hippocampal involvement. It might be the case that hippocampal involvement takes longer to fade. In addition, it has been suggested that hippocampal involvement during memory retrieval processes might never fade completely (Barry & Maguire, 2019; Frankland et al., 2019). This urges us to reconsider whether, even if RP elicits fast consolidation processes, the hippocampus can truly be expected not to play a role in the process of retrieval.

Thirdly, with regards to the ECA, our findings consistently show that RP selectively increases recollection – and not familiarity – which has been predicted by the ECA (Karpicke et al., 2014): Through the update of temporal context, and thus, episodic information at and during *retrieval*, more episodic information is assumed to become available for the next retrieval instance and thus memory accuracy increases. Furthermore, the high pattern similarity in cortical regions (frontal and parietal) lends some support to the ECA on a neural level, as these parts of the neocortex have been linked to episodic memory previously (Buckner et al., 1999; Vilberg & Rugg, 2008).

As an aside, our findings also link to the concept of effort: reduced (low) alpha power points to increased effort in RP as well, and retrieval effort has indeed been discussed as an important factor for RPEs (Pyc & Rawson, 2009). However, retrieval effort or desirable difficulty accounts, which explain how difficult tasks can lead to enhanced performance on future occasions (Bjork & Bjork, 2011) are not unique to RP settings, and are therefore not considered in detail here when it comes to identifying specific processes that underlie RP benefits.

In conclusion, all three of these theories respectively make different predictions, succeed in explaining parts of the RP phenomenon, but often fail in predicting or explaining the totality of RP results. However, all three are focusing on different memory processes and are not mutually exclusive. Based on the findings made in this dissertation, I propose to combine the core concepts of each theory into an integrative framework of RP.

The integrative model I am proposing combines the strengths of all the discussed theoretical accounts of RP into one comprehensive model (see **Figure 6.1**): Firstly, RP is viewed as a marker of importance: the retrieval of a particular item (e.g., a poem) from memory signals

that this item is relevant. Having had the need to retrieve a particular item of information at least once marks this item as relevant and important for future reference, because the necessity for retrieval even once indicates this information might also have potential value again in future. It is logical therefore, that any retrieval occasion could induce mechanisms that lead to the stronger inscription of this information to long-term memory. Thus, because of the relative importance of that item, RP elicits memory processes (encoding, consolidation, retrieval) to be enhanced in order to facilitate future retrieval trials. This process enhancement is largely in line with the previously discussed accounts of RP (see chapter 1.4 Theoretical accounts of the Retrieval Practice Effect. Secondly, beneficial effects of RP on each memory aspect or process are independent, and additive. This means that enhancement of encoding processes might take place alongside an enhancement of consolidation processes, as well as semantic effects alongside episodic effects, and the beneficial effects on memory performance add onto each other, leading to stronger RPEs, the more processes are involved. Thirdly, the model is neither modality, nor memory system specific: Memory content from each sensory modality might benefit from RP, and both declarative and non-declarative memories could become enhanced through RP as it is described in this model.

To further characterize the enhancements of each memory process (encoding, consolidation, and retrieval), as well as semantic and episodic memory, I turn back to the ERH, FRC, and ECA: Based on the current findings, small adjustments need to be made to each theory when implemented into the model. In terms of the ERH, semantic elaboration should not be considered a conscious process, but an automatic effect of RP. As our results show, RP leads to the involvement of semantic processes without the participants being explicitly instructed to perform any semantic elaboration techniques. Even though this account is called

“Elaborative Retrieval”, the processes that are described resemble those seen in elaborative encoding studies (see Karpicke & Blunt, 2011; Karpicke & Smith, 2012; Lehman et al., 2014): in the ERH framework, RP offers a further possibility of elaborating information related to the target, much like an additional encoding opportunity. In terms of integrating the FRC hypothesis into my model, RP can enhance neocortical representations, but does not seem to lead to hippocampal independence, at least not on short time-scales. Whether this is true for longer timescales should be investigated further. However, our results in chapter 5 suggest that consolidation-like processes play a role in RP. In terms of the ECA, the key adjustment is to acknowledge that the processes described as responsible for RP in this account are not sufficient, as they are unable to explain the involvement of semantic processes in the RPE, nor the RP benefits for novel stimuli. The ECA focusses on processes which are part of retrieval: the reinstatement of a context (from initial encoding) is achieved through retrieval, which in turn then leads to an update of the initial context to integrate the new context at retrieval. This account focusses mainly on retrieval, as it describes RP in terms of its effect on future retrieval instances.

The proposed model which integrates the ERH, FRC hypothesis, and ECA, also integrates all three memory processes: Firstly, in terms of encoding, RP can lead to enhanced semantic elaboration. Secondly, RP activates brain areas involved in consolidation and leads to fast enhancement of memory. Lastly, the retrieval itself serves to update the memory context, enhancing episodic memory. To come back to the example of learning a poem by heart, RP enhances memory performance through a combination of semantic elaboration, consolidation mechanisms, and context updates, as seen through increased recollection rates.

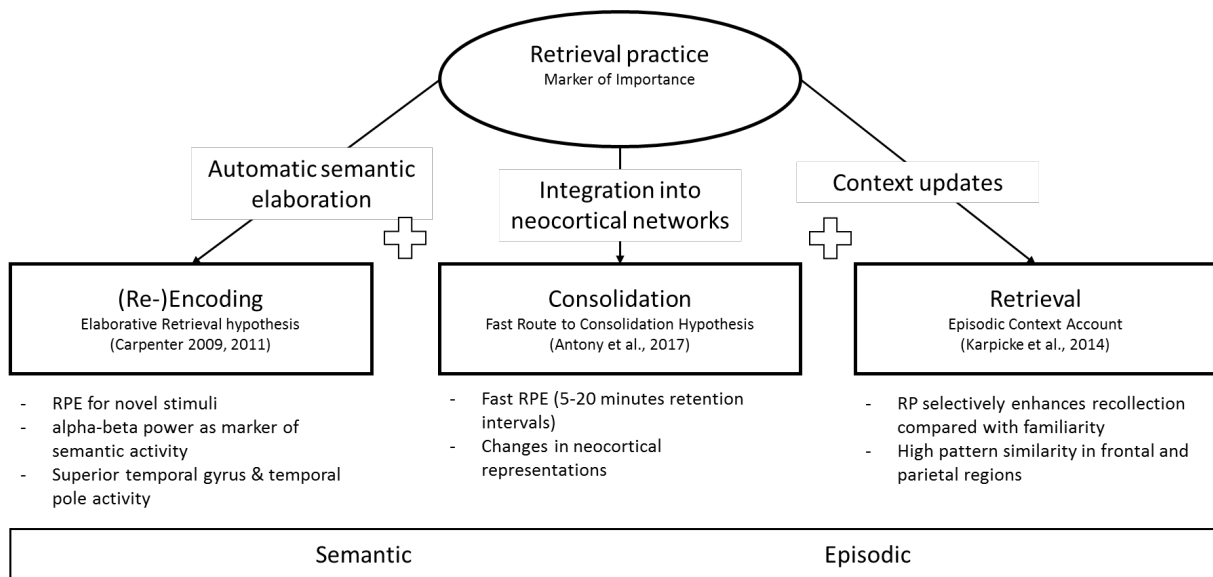


Figure 6.1: Schematic of how RP influences each memory process, as proposed by different explanatory accounts, and our findings pertaining to these theories. All three accounts complement each other to form an integrative view of RP as a learning technique that influences the whole memory-formation process.

This integrative Encoding-Consolidation-Retrieval model of RP can also help shed light on the divergent results as to the effectiveness of RP in old age. Memory generally declines with age, but it does not do so uniformly, neither across participants, nor across memory processes. As such, the strong variability in age-related decreases of memory functioning means that certain underlying mechanisms of RP will be more intact than others, depending on the individual, in turn explaining the lack of consensus as to the degree of RP benefits in old age across the literature. The here proposed model makes clear predictions as to the extent that old individuals still benefit from RP, which should be studied in further experiments: if all three memory processes are preserved, the individual will benefit as much as a comparable individual of younger age. The more of the processes are impaired, the smaller the RP benefit will be.

In addition, as this model offers a modality-general approach to RP, the question arises whether content from different sensory modalities, or even different memory systems, such as procedural memory, could also benefit from RP. Aspects of Encoding, Consolidation, and Retrieval exist in each sensory modality. Therefore, memory content in each sensory modality might benefit from RP. This prediction of the model remains to be tested. Similarly, non-declarative memory also underlies the general memory formation processes of encoding, consolidation and retrieval. The model predicts that RP might span other memory systems, such as procedural memory. So far, some evidence exists in favor of this assumption (Boutin et al., 2013; Tempel & Kubik, 2017) on a behavioral level. It is unclear as of yet, whether RPEs in procedural memory are subject to the same neural underpinnings (see chapters 3 and 5) as declarative memory contents.

Finally, this integrative model of RP influencing encoding, consolidation, and retrieval, can help remove the fallacy of trying to ascertain the superiority of one theory over another from RP research, and thereby redirect the focus of this research to more important questions, such as those outlined below (chapter 6.4).

6.4 Future directions

This dissertation demonstrates that RP can be a powerful tool for improving memory performance in older and younger adults alike. However, a number of further questions have arisen from the research here and still remain to be addressed.

First and foremost, there is the question of the ultimate causes of RP: Why is it that retrieving information leads to its enhancement? Viewing retrieval as a marker of the importance and necessity of a memory item can help answer this question (see chapter 3). This view of

retrieval and retrieval practice is supported by findings showing that RP makes memory less susceptible to interference, and that RP counteracts stress-related memory-performance decreases (Smith et al., 2016). However, the concept of RP as a marker of importance should be explored by trying to manipulate importance of information in other ways and comparing memory performance to RP.

In terms of age-related memory declines, it is highly unlikely that RP can counteract the detrimental effects that aging has on memory, or disorders that affect long-term memory. However, our results suggest that RP could be used as a learning technique to preserve functional levels of memory, and facilitate independent living in older age by providing a tool with which day-to-day demands on memory can be potentially met. Such close-to-life implementations should be investigated, as RP might have the potential to help older individuals cope with the demands they face in everyday life. But, to use RP to improve daily life, it is important to understand its working mechanisms. Furthermore, as suggested in chapter 6.1, the question whether alpha-beta power in relation to RP might be an indicator of brain maintenance could be considered in future studies, as it could offer additional non-invasive information on an individual's central nervous integrity.

In terms of the temporal dynamics of RP, while RP benefits clearly persists over a week-long period, can it last over longer periods of time, which are more relevant for navigating life, such as months and years? Future studies should increase the retention interval between retrieval practice and final recall even further, to span multiple months (Eriksson et al., 2011), to assess the longevity of this remarkable effect.

The potential uses of RP stretch beyond (pre-) clinical samples, into the daily routine of schoolchildren, students, and working adults. Learning new information, or skills continues to

be important throughout the life-span, therefore RP components already employed in some contexts, such as school or university tests and exams, should be expanded upon. Here, it might be beneficial to separate retrieval instances that mainly target learning, from those that are used to assess performance in a competitive context, in order to increase the acceptance of such measures for learning improvement. While RP is a promising avenue to improve memory performance for materials learned through retrieval (Smith et al., 2016), it is also unclear whether training individuals to use a “retrieval mode” consistently to improve their memory performance also translates into improving other aspects of cognitive functioning, and whether it can be used as a measure to ensure a successful academic career for as many children and young adults as possible, as well as improving quality of life in old adulthood.

We have made advances in identifying and understanding the neural bases of the RPE, but further investigation is still necessary. Based on our findings in chapter 3, the question arises whether the observed alpha-beta decrease is instrumental in creating the RPE, or rather a product of semantic involvement and attentional demands. To address this question, methods that directly influence the brain activity, such as transcranial anodal current stimulation (tACS, Antal & Paulus, 2013) could be used to experimentally manipulate alpha-beta power and investigate whether this leads to RP-like behavioral results. Similarly, transcranial magnetic stimulation (TMS, Hallett, 2000) could be employed to ascertain whether the brain activation observed in chapter 5 causally produces RP-like behavioral results, as is suggested by the correlation.

The relevance of neurotransmitters, most importantly Dopamine (DA), should be investigated more directly than has been done here by proxy of novelty seeking. Memory functions are tightly related to DA (Apitz & Bunzeck, 2013; Schott et al., 2006). DA also decreases with age

and its decrease has been linked to memory impairments (Bäckman et al., 2006; Chowdhury et al., 2012). Based on our findings of often reduced RPEs in older participants, it is quite likely that DA might be involved in the processes underlying RP. The same might be the case for Noradrenaline which also plays a key role in memory formation and consolidation (Gibbs et al., 2010), and acetylcholine (Blokland, 1995; Hasselmo, 1999). Understanding the electrophysiological and neurobiological foundations of the RPE in more detail will increase our understanding of memory processes in general. Furthering our understanding of memory must be one of the key goals for future research, especially considering the burden the memory deficits of the increasingly older population (at least in industrialized nations (World Health Organization, 2015) poses on individuals and families, as well as the health care system and the economy.

To achieve a better description of RP and identify further important research questions about the RPE, a general model of this effect needs to be put forward. Firstly, a thorough up-to-date review of the existing literature is necessary. The last meta-analytic review (Rowland, 2014) does not encompass some new findings that have added to our understanding of the RPE. Then, a new model needs to be proposed to the scientific community and thoroughly investigated. The above outlined integrative model could be a such a model, or a starting point from which to elaborate further aspects of RP and make clear predictions as to future findings, e.g. the presence of RP outside of declarative memory, such as in procedural memory (see chapter 6.3), and the exact interplay between retrieval, consolidation, and encoding during retrieval practice. Further, it is also important to investigate whether RP is an effect particular to visual and semantic stimuli, and whether it might also be present in auditory, or olfactory, or other sensory modalities, in order to understand whether the RPE is a modality specific, or

general neural phenomenon (see chapter 6.3). Interestingly, reaction times have not been considered much in the description of the RPE, and might offer insights as to the underlying mechanisms: if RP elaborates a cue-set (see ERH), reaction times might increase as a function of RP-related search-set increases. Under the assumptions of the ECA, however, reaction times should decrease due to restrictions of the search set. If both theories hold true, reaction times might not differ between RP and re-encoded stimuli.

The key processes of memory formation, i.e., encoding, consolidation and retrieval, all seem to be involved in RP. Especially how RP affects consolidation processes should be considered more: markers of neural plasticity, i.e. microstructural changes in the brain could be indicative of consolidation-related processes (Brodt et al., 2018) and might also be present after retrieval practice. Another process of memory formation that has not been considered yet in the RP literature is reconsolidation, which is the phenomenon that memory is made susceptible to change, and enhancement, through retrieval (Schwabe et al., 2014; Walker et al., 2003). The parallels between RP and reconsolidation are obvious, in that RP enhances a memory through the act of retrieval. Identifying the exact communalities between the processes elicited by RP and reconsolidation processes can help in formulating reasonable hypothesis about RP based on the reconsolidation literature, but also further our understanding of general memory function and processes.

On a final note, research on RP is still very heterogeneous, which on the one hand serves to demonstrate its robustness and versatility, but on the other hand, makes it difficult to compare different studies. To better understand how differences in RPE findings arise, direct comparisons of different stimulus materials, memory test format at final recall (cued versus

free-recall versus recognition), feedback availability, and many other factors should be performed.

6.5 Strengths and limitations

Firstly, a core strength of this work lies in the comparative approach between age groups taken to investigate age-related differences in the RPE. Not only is it important to understand mechanisms that can improve memory in older age, where this cognitive function typically exhibits drastic declines. In addition, older participants also typically exhibit higher variance in their responses as well as the neural correlates underlying their responses, which can make uncovering relationships between the brain and behavior more likely (e.g. the relationship between alpha-beta power and RP benefit which surfaced only in older adults). Secondly, our paradigm has reliably led to RPEs across multiple attempts, which makes it interesting for investigating the neural substrates of this effect. The paradigm is characterized by a high number of trials (a total of 160 items in each practice condition in Phase 2), made possible by the use of pictorial stimuli – instead of the often employed word lists. With pictorial stimuli, however, only recognition memory, and not free recall, can be investigated. The study of free recall in the RPE has been important and fruitful, but in order to uncover subtle neural differences between retested and restudied items, a high number of trials is needed for each participant, making word lists less useful. The cognitive load to memorize and reproduce 160 words is more challenging than recognizing 160 images, meaning that word list learning studies have to use fewer trials due to possible floor effects. Thirdly, the paradigm can be easily adapted to small changes in the experimental protocol, such as increasing or decreasing

retention intervals, changing response options at final recall, or using more, or different, stimuli.

Lastly and importantly, this dissertation provides a collection of studies investigating all three processes of memory formation (encoding, consolidation and retrieval) with different neuroimaging techniques and analysis methods (univariate and multivariate procedures, time-frequency analysis), allowing us to ask questions about the psychological and behavioral basis of RP, its temporal neural dynamics, as well as its functional processing inside the brain (fMRI). The comparatively large sample size per study (60-90 participants, amounting to a total sample size of roughly 210 participants), and trial size allowed for strong statistical power for detecting the RPE. Furthermore, we were able to replicate not only our own findings, but also (behavioral) findings from the literature. All three measures of neural and behavioral processing (temporal electrophysiology, metabolic activation and activation patterns, and behavioral output) can help answer different questions about the underlying processes and mechanisms of RP. A further strength of this dissertation lies in the inclusion of these different measures, as well as the successful attempts to link behavior to measures of neural processing, using state-of-the-art analysis methods (RSA, CBP, Bayesian statistics).

When it comes to the limitations of our studies, the use of pictorial stimuli prohibits the investigation of active recall memory (see above). Effects of RP might differ as a function of the final recall test mode, active recall versus passive recognition, and it is not possible to quantify these differences in this paradigm. However, this limitation to recognition memory is compensated by the added possibility of investigating source memory, in terms of recollection versus familiarity judgements, which is not possible in recall studies. In the implementation of “remember-know” paradigms, we gave participants a further choice

option (“guess”). However, this ultimately decreased our statistical power, especially in the fMRI study, making it impossible to consider source memory on a neural level due to insufficient trial numbers per participant. For future studies, it should be considered to not include the “guess” response option, forcing participants to choose between “remember”, “know”, and “new” only. This would in turn lower the difficulty of the final retrieval task, which some older participants reportedly struggled with. The fact that we only used of pictorial stimuli also means that our conclusions about the RPE need to be validated with other types of stimulus material, to ascertain that the here observed results are generalizable. The ecological validity of the results might also be restricted due to the nature of the task. Memory demands in everyday life usually pertain more to planning of errands, remembering names, dates or contact details, as well as ways (Sunderland et al., 1983). This work aimed at laying the ground work for following, more naturalistic tasks and investigations. As of yet, it would be premature to suppose that the effects that are so clearly observable in highly standardized tasks such as ours will have an identical translation into life. Further studies are needed to assert this.

Finally, even though our first two studies were not directly aimed at investigating the veracity of any particular theory, we were able to make many observations pertaining to theories of RP. In summary, the limitations of our study do not meaningfully diminish the conclusions we were able to draw from our data, and the paradigm we employed was particularly suited for investigating subtle neural bases of RP in older and younger participants.

6.6 Conclusion

Retrieval practice is an efficient way to improve memory for specific contents in younger and older adults. The RPE works fast, selectively improves recollection, and also affects novel stimuli. Furthermore, its effects are stable over a period of at least seven days in both young and older adults, making it an interesting candidate to improve memory performance, which can in turn improve quality of life in older age. RP involves semantic as well as episodic memory, and underlying neural markers of the involvement of these processes are linked to behavioral benefits experienced through RP.

Based on different existing accounts of RP, I put forward an integrative model that identifies RP as a learning technique which involves all levels of memory processing, i.e., encoding, consolidation, and retrieval. Future studies should continue investigating the neural bases of the retrieval practice effect, as well as its temporal stability, and translation into real-life situations. This dissertation and its studies contributed to addressing some of the mechanisms behind RP, and its beneficial effects across age.

List of Tables

Table 3.1 Number of averaged responses for Phase 2.	49
Table 3.2 D' values for young and older participants.	50
Table 3.3 Number of averaged responses for Phases 3 and 4.....	50
Table 3.4 Results of the TFA.	55
Table 4.1 Demographic data.....	69
Table 4.2 Reaction times in Phase 2.	73
Table 4.3 Spearman correlation of TCI and RPE	80
Table 5.1 Memory accuracy in Phase 2.	97
Table 5.2 Reaction times in Phase 2.	97
Table 5.3 GLM older > younger	103
Table 5.4 GLM younger > older	105
Table 5.5 GLM new > old	106
Table 5.6 Multiple regression	106
Table 5.7 T-Tests for main effect of ROI in the RSA analysis	111

List of Figures

Figure 2.1 Experimental paradigm.....	29
Figure 3.1 Experimental paradigm.....	46
Figure 3.2 Descriptive results	51
Figure 3.3 Main effects of recognition memory performance in Phases 3 and 4.....	52
Figure 3.4 Age-dependent RPEs.....	52
Figure 3.5 Time frequency results: Main effects	56
Figure 3.6 Time frequency results: Task by age.....	57
Figure 3.7 Correlation analysis	58
Figure 4.1 Experimental paradigm.....	70
Figure 4.2 Retrieval performance for Day 1 and Day 7.....	75
Figure 4.3 Task by age interaction	76
Figure 4.4 Memory by task interaction.....	76
Figure 4.5 Task by novelty by age interaction	79
Figure 5.1 Experimental paradigm.....	91
Figure 5.2 Location of the ROIs for the RSA.....	95
Figure 5.3 Main effects in Phase 3	99
Figure 5.4 Interactions.....	100
Figure 5.5 Age effects: Old > Young	103
Figure 5.6 Age effects: Young > Old.....	104
Figure 5.7 Novelty effects.....	105
Figure 5.8 Multiple regression	107
Figure 5.9 Main effect of task in the RSA.....	109

Figure 5.10 Main effect of ROI in the RSA 109

Figure 5.11 3-way interaction of task, ROI, and age 110

Figure 6.1 Integrative Model 130

Glossary

This glossary contains acronyms introduced in chapters 1, 2 and 6.

CBP: cluster-based permutation

ECA: Episodic Context Account

EEG: electroencephalography

ERH: Elaborative Retrieval Hypothesis

fMRI: functional magnetic resonance imaging

FRC: Fast-Route to Consolidation (Account)

GLM: general linear model

HRF: hemodynamic response function

MTL: medial temporal lobe

PFC: prefrontal cortex

RP: retrieval practice

RPE: retrieval practice effect

RSA: representational similarity analysis

References

- Abbott, E. E. (1909). On the analysis of the factor of recall in the learning process. *The Psychological Review: Monograph Supplements*, 11(1), 159.
- Allen, P. A., Sliwinski, M., Bowie, T., & Madden, D. J. (2002). Differential age effects in semantic and episodic memory. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 57(2), P173–P186.
- Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of verbal learning and verbal behavior*, 22(3), 261–295.
- Anderson, M. C., Bjork, E. L., & Bjork, R. A. (2000). Retrieval-induced forgetting: Evidence for a recall-specific mechanism. *Psychonomic bulletin & review*, 7(3), 522–530.
- Anderson, N. D., Ebert, P. L., Jennings, J. M., Grady, C. L., Cabeza, R., & Graham, S. J. (2008). Recollection-and familiarity-based memory in healthy aging and amnesic mild cognitive impairment. *Neuropsychology*, 22(2), 177.
- Antal, A., & Paulus, W. (2013). Transcranial alternating current stimulation (tACS). *Frontiers in Human Neuroscience*, 7. <https://doi.org/10.3389/fnhum.2013.00317>
- Antony, J. W., Ferreira, C. S., Norman, K. A., & Wimber, M. (2017). Retrieval as a Fast Route to Memory Consolidation. *Trends in Cognitive Sciences*, 21(8), 573–576. <https://doi.org/10.1016/j.tics.2017.05.001>
- Apitz, T., & Bunzeck, N. (2013). Dopamine controls the neural dynamics of memory signals and retrieval accuracy. *Neuropsychopharmacology*, 38(12), 2409–2417.
- Arthurs, O. J., & Boniface, S. (2002). How well do we understand the neural origins of the fMRI BOLD signal? *Trends in Neurosciences*, 25(1), 27–31. [https://doi.org/10.1016/S0166-2236\(00\)01995-0](https://doi.org/10.1016/S0166-2236(00)01995-0)

Astrand, I., Astrand, P. O., Hallbäck, I., & Kilbom, A. (1973). Reduction in maximal oxygen uptake with age. *Journal of Applied Physiology*, *35*(5), 649–654.

Atkinson, R. C., & Shiffrin, R. M. (1968). *Human memory: A proposed system and its control processes*.

Atkinson, R. C., & Shiffrin, R. M. (2016). 23 Human Memory: A Proposed System and Its Control Processes. *Scientists making a difference: One hundred eminent behavioral and brain scientists talk about their most important contributions*, 115.

Backhaus, J., Born, J., Hoeckesfeld, R., Fokuhl, S., Hohagen, F., & Junghanns, K. (2007). Midlife decline in declarative memory consolidation is correlated with a decline in slow wave sleep. *Learning & memory*, *14*(5), 336–341.

Bäckman, L., Nyberg, L., Lindenberger, U., Li, S.-C., & Farde, L. (2006). The correlative triad among aging, dopamine, and cognition: current status and future prospects. *Neuroscience & Biobehavioral Reviews*, *30*(6), 791–807.

Balota, D. A., Duchek, J. M., & Paullin, R. (1989). Age-related differences in the impact of spacing, lag, and retention interval. *Psychology and Aging*, *4*(1), 3.

Bamidis, P. D., Vivas, A. B., Styliadis, C., Frantzidis, C., Klados, M., Schlee, W., Siountas, A., & Papageorgiou, S. G. (2014). A review of physical and cognitive interventions in aging. *Neuroscience & Biobehavioral Reviews*, *44*, 206–220.

Bandettini, P. A., & Ungerleider, L. G. (2001). From neuron to BOLD: new connections. *Nature Neuroscience*, *4*(9), 864–866. <https://doi.org/10.1038/nn0901-864>

Bardo, M. T., Donohew, R. L., & Harrington, N. G. (1996). Psychobiology of novelty seeking and drug seeking behavior. *Behavioural brain research*, *77*(1–2), 23–43.

Barry, D. N., & Maguire, E. A. (2019). Remote memory and the Hippocampus: a constructive critique. *Trends in cognitive sciences*, 23(2), 128–142.

Barulli, D., & Stern, Y. (2013). Efficiency, capacity, compensation, maintenance, plasticity: emerging concepts in cognitive reserve. *Trends in cognitive sciences*, 17(10), 502–509.

Basile, B. M., & Hampton, R. R. (2013). Recognition errors suggest fast familiarity and slow recollection in rhesus monkeys. *Learning & Memory*, 20(8), 431–437.

Benchenane, K., Tiesinga, P. H., & Battaglia, F. P. (2011). Oscillations in the prefrontal cortex: a gateway to memory and attention. *Current opinion in neurobiology*, 21(3), 475–485.

Benjamin, J., Li, L., Patterson, C., Greenberg, B. D., Murphy, D. L., & Hamer, D. H. (1996). Population and familial association between the D4 dopamine receptor gene and measures of novelty seeking. *Nature genetics*, 12(1), 81.

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. *Psychology and the real world: Essays illustrating fundamental contributions to society*, 2(59–68).

Bjork, R. A. (1994). Memory and metamemory considerations in the. *Metacognition: Knowing about knowing*, 185.

Bjork, R. A. (1999). *F 5 assessing our own competence: Heuristics and illusions*.

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

Blokland, A. (1995). Acetylcholine: a neurotransmitter for learning and memory? *Brain Research Reviews*, 21(3), 285–300.

Boehler, C. N., Hopf, J.-M., Krebs, R. M., Stoppel, C. M., Schoenfeld, M. A., Heinze, H.-J., & Noesselt, T. (2011). Task-load-dependent activation of dopaminergic midbrain areas in the absence of reward. *Journal of Neuroscience*, *31*(13), 4955–4961.

Bollmann, S., Puckett, A. M., Cunnington, R., & Barth, M. (2018). Serial correlations in single-subject fMRI with sub-second TR. *NeuroImage*, *166*, 152–166.

Boutin, A., Panzer, S., & Blandin, Y. (2013). Retrieval practice in motor learning. *Human movement science*, *32*(6), 1201–1213.

Bouwmeester, S., & Verkoeijen, P. P. (2011). Why do some children benefit more from testing than others? Gist trace processing to explain the testing effect. *Journal of Memory and Language*, *65*(1), 32–41.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial vision*, *10*(4), 433–436.

Bridge, D. J., & Paller, K. A. (2012). Neural correlates of reactivation and retrieval-induced distortion. *Journal of Neuroscience*, *32*(35), 12144–12151.

Brodts, S., Gais, S., Beck, J., Erb, M., Scheffler, K., & Schönauer, M. (2018). Fast track to the neocortex: A memory engram in the posterior parietal cortex. *Science*, *362*(6418), 1045–1048. <https://doi.org/10.1126/science.aau2528>

Brodts, S., Pöhlchen, D., Flanagin, V. L., Glasauer, S., Gais, S., & Schönauer, M. (2016). Rapid and independent memory formation in the parietal cortex. *Proceedings of the National Academy of Sciences*, *113*(46), 13251–13256.

Buckner, R. L., Kelley, W. M., & Petersen, S. E. (1999). Frontal cortex contributes to human memory formation. *Nature neuroscience*, *2*(4), 311–314.

Bunge, S. A., Burrows, B., & Wagner, A. D. (2004). Prefrontal and hippocampal contributions to visual associative recognition: interactions between cognitive control and episodic retrieval. *Brain and cognition*, *56*(2), 141–152.

Bunzeck, N., & Düzel, E. (2006). Absolute coding of stimulus novelty in the human substantia nigra/VTA. *Neuron*, *51*(3), 369–379.

Bunzeck, N., Guitart-Masip, M., Dolan, R. J., & Düzel, E. (2014). Pharmacological dissociation of novelty responses in the human brain. *Cerebral Cortex*, *24*(5), 1351–1360.

Bunzeck, N., Schütze, H., Stallforth, S., Kaufmann, J., Düzel, S., Heinze, H.-J., & Düzel, E. (2007). Mesolimbic novelty processing in older adults. *Cerebral Cortex*, *17*(12), 2940–2948.

Butler, A. C., & Roediger III, H. L. (2007). Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology*, *19*(4–5), 514–527.

Buxton, R. B. (2001). The elusive initial dip. *Neuroimage*, *13*(6 Pt 1), 953–958.

Cabeza, R. (2002a). Hemispheric asymmetry reduction in older adults: the HAROLD model. *Psychology and aging*, *17*(1), 85.

Cabeza, R. (2002b). Hemispheric asymmetry reduction in older adults: the HAROLD model. *Psychology and aging*, *17*(1), 85.

Cabeza, R., Albert, M., Belleville, S., Craik, F. I., Duarte, A., Grady, C. L., Lindenberger, U., Nyberg, L., Park, D. C., & Reuter-Lorenz, P. A. (2018). Maintenance, reserve and compensation: the cognitive neuroscience of healthy ageing. *Nature Reviews Neuroscience*, *1*.

Cabeza, R., Anderson, N. D., Locantore, J. K., & McIntosh, A. R. (2002). Aging gracefully: compensatory brain activity in high-performing older adults. *Neuroimage*, *17*(3), 1394–1402.

Cabeza, R., Ciaramelli, E., Olson, I. R., & Moscovitch, M. (2008). The parietal cortex and episodic memory: an attentional account. *Nature Reviews Neuroscience*, *9*(8), 613–625.

Cabeza, R., Grady, C. L., Nyberg, L., McIntosh, A. R., Tulving, E., Kapur, S., Jennings, J. M., Houle, S., & Craik, F. I. (1997). Age-related differences in neural activity during memory encoding and retrieval: a positron emission tomography study. *Journal of neuroscience*, *17*(1), 391–400.

Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: the benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(6), 1563.

Carpenter, S. K. (2011). Semantic information activated during retrieval contributes to later retention: Support for the mediator effectiveness hypothesis of the testing effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(6), 1547.

Carpenter, S. K., Pashler, H., & Cepeda, N. J. (2009). Using tests to enhance 8th grade students' retention of US history facts. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *23*(6), 760–771.

Chan, J. C., & McDermott, K. B. (2007). The testing effect in recognition memory: a dual process account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(2), 431.

Chan, J. C., McDermott, K. B., & Roediger III, H. L. (2006). Retrieval-induced facilitation: Initially nontested material can benefit from prior testing of related material. *Journal of Experimental Psychology: General*, *135*(4), 553.

Cho, K. W., Neely, J. H., Crocco, S., & Vitrano, D. (2017). Testing enhances both encoding and retrieval for both tested and untested items. *Quarterly Journal of Experimental Psychology*, *70*(7), 1211–1235. <https://doi.org/10.1080/17470218.2016.1175485>

Chowdhury, R., Guitart-Masip, M., Bunzeck, N., Dolan, R. J., & Duzel, E. (2012). Dopamine Modulates Episodic Memory Persistence in Old Age. *Journal of Neuroscience*, 32(41), 14193–14204. <https://doi.org/10.1523/JNEUROSCI.1278-12.2012>

Ciranni, M. A., & Shimamura, A. P. (1999). Retrieval-induced forgetting in episodic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(6), 1403.

Cloninger, C. R. (1987). A systematic method for clinical description and classification of personality variants: A proposal. *Archives of general psychiatry*, 44(6), 573–588.

Cloninger, C. R., Przybeck, T. R., Svrakic, D. M., & Wetzel, R. D. (1994). *The Temperament and Character Inventory (TCI): A guide to its development and use*.

Coane, J. H. (2013). Retrieval practice and elaborative encoding benefit memory in younger and older adults. *Journal of Applied Research in Memory and Cognition*, 2(2), 95–100.

Craik, F. I., & McDowd, J. M. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13(3), 474.

Crook, T., Bartus, R. T., Ferris, S. H., Whitehouse, P., Cohen, G. D., & Gershon, S. (1986). *Age-associated memory impairment: Proposed diagnostic criteria and measures of clinical change—report of a national institute of mental health work group*.

Cummins, T. D., & Finnigan, S. (2007). Theta power is reduced in healthy cognitive aging. *International Journal of Psychophysiology*, 66(1), 10–17.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods*, 134(1), 9–21.

Desimone, R., Miller, E. K., Chelazzi, L., & Lueschow, A. (1995). *Multiple memory systems in the visual cortex*.

Deuker, L., Olligs, J., Fell, J., Kranz, T. A., Mormann, F., Montag, C., Reuter, M., Elger, C. E., & Axmacher, N. (2013). Memory consolidation by replay of stimulus-specific neural activity. *Journal of Neuroscience*, *33*(49), 19373–19383.

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

Diedenhofen, B., & Musch, J. (2015). cocor: A Comprehensive Solution for the Statistical Comparison of Correlations. *PLOS ONE*, *10*(4), e0121945. <https://doi.org/10.1371/journal.pone.0121945>

Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience*, *11*(2), 114.

Dimsdale-Zucker, H. R., & Ranganath, C. (2018). Representational similarity analyses: A practical guide for functional MRI applications. In *Handbook of Behavioral Neuroscience* (Bd. 28, S. 509–525). Elsevier.

Dolan, R. J., & Fletcher, P. C. (1997). Dissociating prefrontal and hippocampal function in episodic memory encoding. *Nature*, *388*(6642), 582–585. <https://doi.org/10.1038/41561>

Dolcos, F., Rice, H. J., & Cabeza, R. (2002). Hemispheric asymmetry and aging: right hemisphere decline or asymmetry reduction. *Neuroscience & Biobehavioral Reviews*, *26*(7), 819–825.

Duarte, A., Ranganath, C., Trujillo, C., & Knight, R. T. (2006). Intact recollection memory in high-performing older adults: ERP and behavioral evidence. *Journal of Cognitive Neuroscience*, *18*(1), 33–47.

Düzel, E., Bunzeck, N., Guitart-Masip, M., & Düzel, S. (2010). NOvelty-related motivation of anticipation and exploration by dopamine (NOMAD): implications for healthy aging. *Neuroscience & Biobehavioral Reviews*, *34*(5), 660–669.

Düzel, E., Penny, W. D., & Burgess, N. (2010). Brain oscillations and memory. *Current opinion in neurobiology*, *20*(2), 143–149.

Ebstein, R. P., Novick, O., Umansky, R., Priel, B., Osher, Y., Blaine, D., Bennett, E. R., Nemanov, L., Katz, M., & Belmaker, R. H. (1996). Dopamine D4 receptor (D4DR) exon III polymorphism associated with the human personality trait of novelty seeking. *Nature genetics*, *12*(1), 78.

Eckart, C., Fuentemilla, L., Bauch, E. M., & Bunzeck, N. (2014). Dopaminergic stimulation facilitates working memory and differentially affects prefrontal low theta oscillations. *Neuroimage*, *94*, 185–192.

Engel, A. K., & Fries, P. (2010). Beta-band oscillations—signalling the status quo? *Current opinion in neurobiology*, *20*(2), 156–165.

Eriksson, J., Kalpouzos, G., & Nyberg, L. (2011). Rewiring the brain with repeated retrieval: A parametric fMRI study of the testing effect. *Neuroscience Letters*, *505*(1), 36–40. <https://doi.org/10.1016/j.neulet.2011.08.061>

Euston, D. R., Tatsuno, M., & McNaughton, B. L. (2007). Fast-forward playback of recent memory sequences in prefrontal cortex during sleep. *science*, *318*(5853), 1147–1150.

Fearnley, J. M., & Lees, A. J. (1991). Ageing And Parkinson's Disease: Substantia Nigra Regional Selectivity. *Brain*, *114*(5), 2283–2301. <https://doi.org/10.1093/brain/114.5.2283>

Ferreira, C. S., Charest, I., & Wimber, M. (2019). Retrieval aids the creation of a generalised memory trace and strengthens episode-unique information. *NeuroImage*, *201*, 115996.

Ferrucci, L., Gonzalez-Freire, M., Fabbri, E., Simonsick, E., Tanaka, T., Moore, Z., Salimi, S., Sierra, F., & Cabo, R. de. (2020). Measuring biological aging in humans: A quest. *Aging Cell*, *19*(2), e13080. <https://doi.org/10.1111/accel.13080>

Fink, A., Grabner, R. H., Neuper, C., & Neubauer, A. C. (2005). EEG alpha band dissociation with increasing task demands. *Cognitive brain research*, *24*(2), 252–259.

Fischl, B. (2012). FreeSurfer. *Neuroimage*, *62*(2), 774–781.

Fogel, S. M., Albouy, G., Vien, C., Popovicci, R., King, B. R., Hoge, R., Jbabdi, S., Benali, H., Karni, A., & Maquet, P. (2014). fMRI and sleep correlates of the age-related impairment in motor memory consolidation. *Human brain mapping*, *35*(8), 3625–3645.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”: a practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric research*, *12*(3), 189–198.

Fraga, M. C., de Moura, E. G., da Silva Lima, N., Lisboa, P. C., de Oliveira, E., Silva, J. O., Claudio-Neto, S., Filgueiras, C. C., Abreu-Villaça, Y., & Manhães, A. C. (2014). Anxiety-like, novelty-seeking and memory/learning behavioral traits in male Wistar rats submitted to early weaning. *Physiology & Behavior*, *124*, 100–106. <https://doi.org/10.1016/j.physbeh.2013.11.001>

Frankland, P. W., & Bontempi, B. (2005). The organization of recent and remote memories. *Nature Reviews Neuroscience*, *6*(2), 119–130.

Frankland, P. W., Josselyn, S. A., & Köhler, S. (2019). The neurobiological foundation of memory retrieval. *Nature Neuroscience*, *22*(10), 1576–1585. <https://doi.org/10.1038/s41593-019-0493-1>

Freitas, S., Simões, M. R., Alves, L., & Santana, I. (2013). Montreal cognitive assessment: validation study for mild cognitive impairment and Alzheimer disease. *Alzheimer Disease and Associated Disorders*, *27*(1), 37–43. <https://doi.org/10.1097/WAD.0b013e3182420bfe>

Frey, S., & Petrides, M. (2000). Orbitofrontal cortex: A key prefrontal region for encoding information. *Proceedings of the National Academy of Sciences*, *97*(15), 8723–8727.

Fries, P. (2015). Rhythms for cognition: communication through coherence. *Neuron*, *88*(1), 220–235.

Fritz, C. O., Morris, P. E., Nolan, D., & Singleton, J. (2007). Expanding retrieval practice: An effective aid to preschool children's learning. *The Quarterly Journal of Experimental Psychology*, *60*(7), 991–1004.

Gabrieli, J. D. (1996). Memory systems analyses of mnemonic disorders in aging and age-related diseases. *Proceedings of the National Academy of Sciences*, *93*(24), 13534–13540.

Gabrieli, J. D. (1998). Cognitive Neuroscience of Human Memory. *Annual Review of Psychology*, *49*(1), 87–115. <https://doi.org/10.1146/annurev.psych.49.1.87>

Gao, C., Rosburg, T., Hou, M., Li, B., Xiao, X., & Guo, C. (2016). The role of retrieval mode and retrieval orientation in retrieval practice: insights from comparing recognition memory testing formats and restudying. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(6), 977–990.

Garcia, D., Lester, N., Cloninger, K. M., & Robert Cloninger, C. (2017). Temperament and Character Inventory (TCI). *Encyclopedia of personality and individual differences*, 1–3.

Gazzaniga, M. S. (2000). *The new cognitive neurosciences*. MIT press.

Gibbs, M. E., Hutchinson, D. S., & Summers, R. J. (2010). Noradrenaline release in the locus coeruleus modulates memory formation and consolidation; roles for α - and β -adrenergic receptors. *Neuroscience*, 170(4), 1209–1222.

Gläscher, J., & Gitelman, D. (2008). Contrast weights in flexible factorial design with multiple groups of subjects. *SPM@ JISCMAIL. AC. UK) Sml, editor*, 1–12.

Gola, M., Kamiński, J., Brzezicka, A., & Wróbel, A. (2012). Beta band oscillations as a correlate of alertness—changes in aging. *International Journal of Psychophysiology*, 85(1), 62–67.

Gonzalez, C., Kramar, C., Garagoli, F., Rossato, J. I., Weisstaub, N., Cammarota, M., & Medina, J. H. (2013). Medial prefrontal cortex is a crucial node of a rapid learning system that retrieves recent and remote memories. *Neurobiology of learning and memory*, 103, 19–25.

Goossens, N. A. M. C., Camp, G., Verkoeijen, P. P. J. L., Tabbers, H. K., Bouwmeester, S., & Zwaan, R. A. (2016). Distributed Practice and Retrieval Practice in Primary School Vocabulary Learning: A Multi-classroom Study. *Applied Cognitive Psychology*, 30(5), 700–712. <https://doi.org/10.1002/acp.3245>

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association*, 1(1), 114–125.

Guran, C.-N. A., Herweg, N. A., & Bunzeck, N. (2019). Age-related decreases in the retrieval practice effect directly relate to changes in alpha-beta oscillations. *Journal of Neuroscience*, 39(22), 4344–4352.

Guran, C.-N. A., Lehmann-Grube, J. K., & Bunzeck, N. (2020). Retrieval practice improves recollection-based memory over a seven-day period in younger and older adults. *Frontiers in Psychology, 10*, 2997.

Haas, L. F. (2003). Hans berger (1873–1941), richard caton (1842–1926), and electroencephalography. *Journal of Neurology, Neurosurgery & Psychiatry, 74*(1), 9–9.

Hallett, M. (2000). Transcranial magnetic stimulation and the human brain. *Nature, 406*(6792), 147–150.

Hanslmayr, S., & Staudigl, T. (2014). How brain oscillations form memories—a processing based perspective on oscillatory subsequent memory effects. *Neuroimage, 85*, 648–655.

Hanslmayr, S., Staudigl, T., & Fellner, M.-C. (2012). Oscillatory power decreases and long-term memory: the information via desynchronization hypothesis. *Frontiers in Human Neuroscience, 6*. <https://doi.org/10.3389/fnhum.2012.00074>

Harris, J. A., Petersen, R. S., & Diamond, M. E. (2001). The cortical distribution of sensory memories. *Neuron, 30*(2), 315–318.

Hasselmo, M. E. (1999). Neuromodulation: acetylcholine and memory consolidation. *Trends in cognitive sciences, 3*(9), 351–359.

Haxby, J. V., Connolly, A. C., & Guntupalli, J. S. (2014). Decoding neural representational spaces using multivariate pattern analysis. *Annual review of neuroscience, 37*, 435–456.

Hedden, T., & Gabrieli, J. D. (2004). Insights into the ageing mind: a view from cognitive neuroscience. *Nature reviews neuroscience, 5*(2), 87.

Henkel, L. A. (2007). The benefits and costs of repeated memory tests for young and older adults. *Psychology and Aging, 22*(3), 580.

Herweg, N. A., Solomon, E. A., & Kahana, M. J. (2020). Theta Oscillations in Human Memory. *Trends in Cognitive Sciences*, 24(3), 208–227.
<https://doi.org/10.1016/j.tics.2019.12.006>

Herweg, N. A., Sommer, T., & Bunzeck, N. (2018). Retrieval demands adaptively change striatal old/new signals and boost subsequent long-term memory. *Journal of Neuroscience*, 38(3), 745–754.

Homma, S., Nakajima, Y., Musha, T., Okamoto, Y., & He, B. (1987). Dipole-tracing method applied to human brain potentials. *Journal of neuroscience methods*, 21(2–4), 195–200.

Hoops, S., Nazem, S., Siderowf, A. D., Duda, J. E., Xie, S. X., Stern, M. B., & Weintraub, D. (2009). Validity of the MoCA and MMSE in the detection of MCI and dementia in Parkinson disease. *Neurology*, 73(21), 1738–1745.

Hsieh, L.-T., & Ranganath, C. (2014). Frontal midline theta oscillations during working memory maintenance and episodic encoding and retrieval. *Neuroimage*, 85, 721–729.

Hupbach, A. (2015). Retrieval practice does not safeguard memories from interference-based forgetting. *Learning and Motivation*, 49, 23–30.

Jackson, A. F., & Bolger, D. J. (2014). The neurophysiological bases of EEG and EEG measurement: A review for the rest of us. *Psychophysiology*, 51(11), 1061–1071.
<https://doi.org/10.1111/psyp.12283>

Jacoby, L. L., Shimizu, Y., Velanova, K., & Rhodes, M. G. (2005). Age differences in depth of retrieval: Memory for foils. *Journal of Memory and Language*, 52(4), 493–504.
<https://doi.org/10.1016/j.jml.2005.01.007>

Jensen, O., & Lisman, J. E. (2005). Hippocampal sequence-encoding driven by a cortical multi-item working memory buffer. *Trends in neurosciences*, *28*(2), 67–72.

Ji, D., & Wilson, M. A. (2007). Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nature neuroscience*, *10*(1), 100–107.

Jonker, T. R., Dimsdale-Zucker, H., Ritchey, M., Clarke, A., & Ranganath, C. (2018). Neural reactivation in parietal cortex enhances memory for episodically linked information. *Proceedings of the National Academy of Sciences*, *115*(43), 11084–11089.

JudgeRoy, J. O., Davis III, B., & Öunpuu, S. (1996). Step length reductions in advanced age: the role of ankle and hip kinetics. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *51*(6), M303–M312.

Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, *331*(6018), 772–775.

Karpicke, J. D., Blunt, J. R., & Smith, M. A. (2016). Retrieval-Based Learning: Positive Effects of Retrieval Practice in Elementary School Children. *Frontiers in Psychology*, *7*.
<https://doi.org/10.3389/fpsyg.2016.00350>

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

Karpicke, J. D., Lehman, M., & Aue, W. R. (2014). Retrieval-based learning: An episodic context account. *Psychology of learning and motivation*, *61*, 237–284.

Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science*, *319*(5865), 966–968.

Karpicke, J. D., & Smith, M. A. (2012). Separate mnemonic effects of retrieval practice and elaborative encoding. *Journal of Memory and Language*, *67*(1), 17–29.

Karrasch, M., Laine, M., Rapinoja, P., & Krause, C. M. (2004). Effects of normal aging on event-related desynchronization/synchronization during a memory task in humans. *Neuroscience letters*, *366*(1), 18–23.

Kazantseva, A. V., Gaysina, D. A., Malykh, S. B., & Khusnutdinova, E. K. (2009). Role of dopamine transporter gene (DAT1) polymorphisms in personality traits variation. *Russian journal of genetics*, *45*(8), 974–980.

Keresztes, A., Kaiser, D., Kovács, G., & Racsmány, M. (2013). Testing promotes long-term learning via stabilizing activation patterns in a large network of brain areas. *Cerebral Cortex*, *24*(11), 3025–3035.

Kim, S.-G., Richter, W., & Uğurbil, K. (1997). Limitations of temporal resolution in functional MRI. *Magnetic resonance in medicine*, *37*(4), 631–636.

Kliegl, O., & Bäuml, K.-H. T. (2016). Retrieval practice can insulate items against intralist interference: Evidence from the list-length effect, output interference, and retrieval-induced forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(2), 202.

Klimesch, W. (1996). Memory processes, brain oscillations and EEG synchronization. *International Journal of Psychophysiology*, *24*(1), 61–100. [https://doi.org/10.1016/S0167-8760\(96\)00057-8](https://doi.org/10.1016/S0167-8760(96)00057-8)

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain research reviews*, *29*(2), 169–195.

Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in cognitive sciences*, *16*(12), 606–617.

Klimesch, W., Doppelmayr, M., Russegger, H., Pachinger, T., & Schwaiger, J. (1998). Induced alpha band power changes in the human EEG and attention. *Neuroscience letters*, *244*(2), 73–76.

Klimesch, W., Doppelmayr, M., Schimke, H., & Ripper, B. (1997). Theta synchronization and alpha desynchronization in a memory task. *Psychophysiology*, *34*(2), 169–176.

Klimesch, W., Doppelmayr, M., Stadler, W., Pöllhuber, D., Sauseng, P., & Röhme, D. (2001). Episodic retrieval is reflected by a process specific increase in human electroencephalographic theta activity. *Neuroscience Letters*, *302*(1), 49–52. [https://doi.org/10.1016/S0304-3940\(01\)01656-1](https://doi.org/10.1016/S0304-3940(01)01656-1)

Koen, J. D., & Yonelinas, A. P. (2014). The Effects of Healthy Aging, Amnesic Mild Cognitive Impairment, and Alzheimer's Disease on Recollection and Familiarity: A Meta-Analytic Review. *Neuropsychology review*, *24*(3), 332–354. <https://doi.org/10.1007/s11065-014-9266-5>

Kopell, N., Ermentrout, G. B., Whittington, M. A., & Traub, R. D. (2000). Gamma rhythms and beta rhythms have different synchronization properties. *Proceedings of the National Academy of Sciences*, *97*(4), 1867–1872.

Kornell, N., Bjork, R. A., & Garcia, M. A. (2011). Why tests appear to prevent forgetting: A distribution-based bifurcation model. *Journal of Memory and Language*, *65*(2), 85–97.

Krebs, R. M., Schott, B. H., & Düzel, E. (2009). Personality traits are differentially associated with patterns of reward and novelty processing in the human substantia nigra/ventral tegmental area. *Biological psychiatry*, *65*(2), 103–110.

Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008). Representational similarity analysis-connecting the branches of systems neuroscience. *Frontiers in systems neuroscience*, 2, 4.

Kromann, C. B., Bohnstedt, C., Jensen, M. L., & Ringsted, C. (2010). The testing effect on skills learning might last 6 months. *Advances in Health Sciences Education*, 15(3), 395–401. <https://doi.org/10.1007/s10459-009-9207-x>

Kumaran, D., & Maguire, E. A. (2009). Novelty signals: a window into hippocampal information processing. *Trends in cognitive sciences*, 13(2), 47–54.

Lee, H., Samide, R., Richter, F. R., & Kuhl, B. A. (2019). Decomposing parietal memory reactivation to predict consequences of remembering. *Cerebral cortex*, 29(8), 3305–3318.

Lehman, M., Smith, M. A., & Karpicke, J. D. (2014). Toward an episodic context account of retrieval-based learning: Dissociating retrieval practice and elaboration. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(6), 1787.

Lewis, C. S. (2003). *Out of the silent planet*. New York: Scribner.

Lindenberger, U., & Baltes, P. B. (1994). Sensory functioning and intelligence in old age: a strong connection. *Psychology and aging*, 9(3), 339.

Lisman, J. E., & Grace, A. A. (2005). The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron*, 46(5), 703–713.

Lisman, J., Grace, A. A., & Duzel, E. (2011). A neoHebbian framework for episodic memory; role of dopamine-dependent late LTP. *Trends in neurosciences*, 34(10), 536–547.

Liu, Y., Rosburg, T., Gao, C., Weber, C., & Guo, C. (2017). Differentiation of subsequent memory effects between retrieval practice and elaborative study. *Biological psychology*, 127, 134–147.

Logan, J. M., Sanders, A. L., Snyder, A. Z., Morris, J. C., & Buckner, R. L. (2002). Under-recruitment and nonselective recruitment: dissociable neural mechanisms associated with aging. *Neuron*, *33*(5), 827–840.

Logothetis, N. K. (2003). The Underpinnings of the BOLD Functional Magnetic Resonance Imaging Signal. *The Journal of Neuroscience*, *23*(10), 3963–3971. <https://doi.org/10.1523/JNEUROSCI.23-10-03963.2003>

Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature*, *453*(7197), 869–878.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*(6843), 150–157.

Logothetis, N. K., & Pfeuffer, J. (2004). On the nature of the BOLD fMRI contrast mechanism. *Magnetic resonance imaging*, *22*(10), 1517–1531.

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, J., Ly, A., Gronau, Q. F., Smira, M., & Epskamp, S. (2019). JASP: graphical statistical software for common statistical designs. *Journal of Statistical Software*, *88*(2).

Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT press.

Maillet, D., & Rajah, M. N. (2014). Age-related differences in brain activity in the subsequent memory paradigm: A meta-analysis. *Neuroscience & Biobehavioral Reviews*, *45*, 246–257. <https://doi.org/10.1016/j.neubiorev.2014.06.006>

Mander, B. A., Rao, V., Lu, B., Saletin, J. M., Lindquist, J. R., Ancoli-Israel, S., Jagust, W., & Walker, M. P. (2013). Prefrontal atrophy, disrupted NREM slow waves and impaired hippocampal-dependent memory in aging. *Nature neuroscience*, *16*(3), 357.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of neuroscience methods*, 164(1), 177–190.

Markowitsch, H. J. (2003). Psychogenic amnesia. *Neuroimage*, 20, S132–S138.

Marshall, L., Helgadóttir, H., Mölle, M., & Born, J. (2006). Boosting slow oscillations during sleep potentiates memory. *Nature*, 444(7119), 610–613.

McCourt, W. F., Gurrera, R. J., & Cutter, H. S. (1993). Sensation seeking and novelty seeking. *The Journal of nervous and mental disease*, 181(5), 309–312.

McGaugh, J. L. (2000). Memory—a century of consolidation. *Science*, 287(5451), 248–251.

Meyer, A. N., & Logan, J. M. (2013). Taking the testing effect beyond the college freshman: Benefits for lifelong learning. *Psychology and aging*, 28(1), 142.

Mognon, A., Jovicich, J., Bruzzone, L., & Buiatti, M. (2011). ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology*, 48(2), 229–240.

Morcom, A. M., Good, C. D., Frackowiak, R. S., & Rugg, M. D. (2003). Age effects on the neural correlates of successful memory encoding. *Brain*, 126(1), 213–229.

Morcom, A. M., & Henson, R. N. (2018). Increased prefrontal activity with aging reflects nonspecific neural responses rather than compensation. *Journal of Neuroscience*, 38(33), 7303–7313.

Moreira, B. F. T., Pinto, T. S. S., Starling, D. S. V., & Jaeger, A. (2019). Retrieval Practice in Classroom Settings: A Review of Applied Research. *Frontiers in Education*, 4. <https://doi.org/10.3389/feduc.2019.00005>

Morel, P. (2018). Gramm: grammar of graphics plotting in Matlab. *Journal of Open Source Software*, 3(23), 568.

Moscovitch, M., & Winocur, G. (2002). The frontal cortex and working with memory. *Principles of frontal lobe function*, 188, 209.

Nadel, L., & Moscovitch, M. (1997). Memory consolidation, retrograde amnesia and the hippocampal complex. *Current opinion in neurobiology*, 7(2), 217–227.

Nakamura, K., & Kubota, K. (1995). Mnemonic firing of neurons in the monkey temporal pole during a visual recognition memory task. *Journal of neurophysiology*, 74(1), 162–178.

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695–699.

Nehamas, A., & Woodruff, P. (1995). *Phaedrus*. Hackett Publishing.

Newcombe, N. S., Lloyd, M. E., & Ratliff, K. R. (2007). Development of episodic and autobiographical memory: A cognitive neuroscience perspective. *Advances in child development and behaviour*, 35, 40–89.

Nyberg, L., Backman, L., Erngrund, K., Olofsson, U., & Nilsson, L.-G. (1996). Age Differences in Episodic Memory, Semantic Memory, and Priming: Relationships to Demographic, Intellectual, and Biological Factors. *Education*, 13, 2–60.

Nyberg, L., Habib, R., McIntosh, A. R., & Tulving, E. (2000). Reactivation of encoding-related brain activity during memory retrieval. *Proceedings of the National Academy of Sciences*, 97(20), 11120–11124.

Nyberg, L., Lövdén, M., Riklund, K., Lindenberger, U., & Bäckman, L. (2012). Memory aging and brain maintenance. *Trends in cognitive sciences*, *16*(5), 292–305.

Oberauer, K., & Lewandowsky, S. (2008). Forgetting in immediate serial recall: Decay, temporal distinctiveness, or interference? *Psychological review*, *115*(3), 544.

Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences*, *87*(24), 9868–9872. <https://doi.org/10.1073/pnas.87.24.9868>

Olejniczak, P. (2006). Neurophysiologic basis of EEG. *Journal of clinical neurophysiology*, *23*(3), 186–189.

Olszowy, W., Aston, J., Rua, C., & Williams, G. B. (2019). Accurate autocorrelation modeling substantially improves fMRI reliability. *Nature communications*, *10*(1), 1–11.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational intelligence and neuroscience*, *2011*, 9.

Osipova, D., Takashima, A., Oostenveld, R., Fernández, G., Maris, E., & Jensen, O. (2006). Theta and gamma oscillations predict encoding and retrieval of declarative memory. *Journal of neuroscience*, *26*(28), 7523–7531.

Pastötter, B., & Bäuml, K.-H. T. (2014). Retrieval practice enhances new learning: the forward effect of testing. *Frontiers in Psychology*, *5*. <https://doi.org/10.3389/fpsyg.2014.00286>

Pernet, C. R., Latinus, M., Nichols, T. E., & Rousselet, G. A. (2015). Cluster-based computational methods for mass univariate analyses of event-related brain potentials/fields: A simulation study. *Journal of Neuroscience Methods*, 250, 85–93. <https://doi.org/10.1016/j.jneumeth.2014.08.003>

Petrides, M., Alivisatos, B., Meyer, E., & Evans, A. C. (1993). Functional activation of the human frontal cortex during the performance of verbal working memory tasks. *Proceedings of the National Academy of Sciences*, 90(3), 878–882.

Poldrack, R. A., Mumford, J. A., & Nichols, T. E. (2011). *Handbook of functional MRI data analysis*. Cambridge University Press.

Pyc, M. A., & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? *Journal of Memory and Language*, 60(4), 437–447.

Pyc, M. A., & Rawson, K. A. (2010). Why testing improves memory: Mediator effectiveness hypothesis. *Science*, 330(6002), 335–335.

R Core Team (2013). *R: A language and environment for statistical computing*.

Rabinowitz, J. C., & Craik, F. I. (1986). Prior retrieval effects in young and old adults. *Journal of Gerontology*, 41(3), 368–375.

Rawson, K. A., & Dunlosky, J. (2011). Optimizing schedules of retrieval practice for durable and efficient learning: How much is enough? *Journal of Experimental Psychology: General*, 140(3), 283.

Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., Dahle, C., Gerstorf, D., & Acker, J. D. (2005). Regional Brain Changes in Aging Healthy Adults: General Trends, Individual Differences and Modifiers. *Cerebral Cortex*, *15*(11), 1676–1689. <https://doi.org/10.1093/cercor/bhi044>

Reed, A. V. (1973). Speed-accuracy trade-off in recognition memory. *Science*, *181*(4099), 574–576.

Reuter-Lorenz, P. A., & Park, D. C. (2010). Human neuroscience and the aging mind: a new look at old problems. *The Journals of Gerontology: Series B*, *65*(4), 405–415.

Roediger III, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in cognitive sciences*, *15*(1), 20–27.

Roediger III, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological science*, *17*(3), 249–255.

Rogalski, Y., Altmann, L. J., & Rosenbek, J. C. (2014). Retrieval practice and testing improve memory in older adults. *Aphasiology*, *28*(4), 381–400.

Rosano, C., Simonsick, E. M., Harris, T. B., Kritchevsky, S. B., Brach, J., Visser, M., Yaffe, K., & Newman, A. B. (2005). Association between physical and cognitive function in healthy elderly: the health, aging and body composition study. *Neuroepidemiology*, *24*(1–2), 8–14.

Rosburg, T., Johansson, M., Weigl, M., & Mecklinger, A. (2015). How does testing affect retrieval-related processes? An event-related potential (ERP) study on the short-term effects of repeated retrieval. *Cognitive, Affective, & Behavioral Neuroscience*, *15*(1), 195–210.

Rowland, C. A. (2014). *The effect of testing versus restudy on retention: a meta-analytic review of the testing effect*. American Psychological Association.

Rowland, C. A., & DeLosh, E. L. (2015). Mnemonic benefits of retrieval practice at short retention intervals. *Memory*, 23(3), 403–419. <https://doi.org/10.1080/09658211.2014.889710>

Rugg, M. D., & Vilberg, K. L. (2013). Brain Networks Underlying Episodic Memory Retrieval. *Current opinion in neurobiology*, 23(2), 255–260. <https://doi.org/10.1016/j.conb.2012.11.005>

Sara, S. J., Dyon-Laurent, C., & Hervé, A. (1995). Novelty seeking behavior in the rat is dependent upon the integrity of the noradrenergic system. *Cognitive Brain Research*, 2(3), 181–187.

Schacter, D. L., & Tulving, E. (1994). *Memory systems*. Cambridge, MA: MIT Press.

Schaie, K. W. (1994). The course of adult intellectual development. *American Psychologist*, 49(4), 304–313. <https://doi.org/10.1037/0003-066X.49.4.304>

Schinka, J. A., Letsch, E. A., & Crawford, F. C. (2002). DRD4 and novelty seeking: Results of meta-analyses. *American journal of medical genetics*, 114(6), 643–648.

Schott, B. H., Seidenbecher, C. I., Fenker, D. B., Lauer, C. J., Bunzeck, N., Bernstein, H.-G., Tischmeyer, W., Gundelfinger, E. D., Heinze, H.-J., & Düzel, E. (2006). The dopaminergic midbrain participates in human episodic memory formation: evidence from genetic imaging. *Journal of Neuroscience*, 26(5), 1407–1417.

Schwabe, L., Nader, K., & Pruessner, J. C. (2014). Reconsolidation of Human Memory: Brain Mechanisms and Clinical Relevance. *Biological Psychiatry*, 76(4), 274–280. <https://doi.org/10.1016/j.biopsych.2014.03.008>

Sebastián, M., Reales, J. M., & Ballesteros, S. (2011). Ageing affects event-related potentials and brain oscillations: a behavioral and electrophysiological study using a haptic recognition memory task. *Neuropsychologia*, *49*(14), 3967–3980.

Sharon, T., Moscovitch, M., & Gilboa, A. (2011). Rapid neocortical acquisition of long-term arbitrary associations independent of the hippocampus. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(3), 1146–1151.
<https://doi.org/10.1073/pnas.1005238108>

Small, S. A. (2001). Age-Related Memory Decline: Current Concepts and Future Directions. *Archives of Neurology*, *58*(3), 360–364.
<https://doi.org/10.1001/archneur.58.3.360>

Smith, A. M., Floerke, V. A., & Thomas, A. K. (2016). Retrieval practice protects memory against acute stress. *Science*, *354*(6315), 1046–1048.

Spencer, R. M., Gouw, A. M., & Ivry, R. B. (2007). Age-related decline of sleep-dependent consolidation. *Learning & Memory*, *14*(7), 480–484.

Spencer, W. D., & Raz, N. (1995). Differential effects of aging on memory for content and context: a meta-analysis. *Psychology and aging*, *10*(4), 527.

Spitzer, B., & Haegens, S. (2017). Beyond the status Quo: a role for beta oscillations in endogenous content (Re-) activation. *eneuro*, ENEURO–0170.

Spitzer, H. F. (1939). Studies in retention. *Journal of Educational Psychology*, *30*(9), 641.

Squire, L. R. (1992). Declarative and Nondeclarative Memory: Multiple Brain Systems Supporting Learning and Memory. *Journal of Cognitive Neuroscience*, *4*(3), 232–243.
<https://doi.org/10.1162/jocn.1992.4.3.232>

Squire, L. R., Genzel, L., Wixted, J. T., & Morris, R. G. (2015). Memory consolidation. *Cold Spring Harbor perspectives in biology*, 7(8), a021766.

Squire, L. R., & Wixted, J. T. (2011). The cognitive neuroscience of human memory since HM. *Annual review of neuroscience*, 34, 259–288.

Squire, L. R., & Zola-Morgan, S. (1991). The medial temporal lobe memory system. *Science*, 253(5026), 1380–1386.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior research methods, instruments, & computers*, 31(1), 137–149.

Staresina, B. P., Gray, J. C., & Davachi, L. (2009). Event congruency enhances episodic memory encoding through semantic elaboration and relational binding. *Cerebral Cortex*, 19(5), 1198–1207.

Stickgold, R. (2005). Sleep-dependent memory consolidation. *Nature*, 437(7063), 1272.

Stippich, C., Ochmann, H., & Sartor, K. (2002). Somatotopic mapping of the human primary sensorimotor cortex during motor imagery and motor execution by functional magnetic resonance imaging. *Neuroscience letters*, 331(1), 50–54.

St-Laurent, M., Abdi, H., Burianová, H., & Grady, C. L. (2011). Influence of Aging on the Neural Correlates of Autobiographical, Episodic and Semantic Memory Retrieval. *Journal of cognitive neuroscience*, 23(12), 4150–4163. https://doi.org/10.1162/jocn_a_00079

Sumowski, J. F., Chiaravalloti, N., & DeLuca, J. (2010). Retrieval practice improves memory in multiple sclerosis: Clinical application of the testing effect. *Neuropsychology*, 24(2), 267.

Sumowski, J. F., Coyne, J., Cohen, A., & DeLuca, J. (2014). Retrieval practice improves memory in survivors of severe traumatic brain injury. *Archives of physical medicine and rehabilitation, 95*(2), 397–400.

Sunderland, A., Harris, J. E., & Baddeley, A. D. (1983). Do laboratory tests predict everyday memory? A neuropsychological study. *Journal of verbal learning and verbal behavior, 22*(3), 341–357.

Tempel, T., & Kubik, V. (2017). Test-potentiated learning of motor sequences. *Memory, 25*(3), 326–334. <https://doi.org/10.1080/09658211.2016.1171880>

Tombaugh, T. N., & McIntyre, N. J. (1992). The mini-mental state examination: a comprehensive review. *Journal of the American Geriatrics Society, 40*(9), 922–935.

Tomita, H., Ohbayashi, M., Nakahara, K., Hasegawa, I., & Miyashita, Y. (1999). Top-down signal from prefrontal cortex in executive control of memory retrieval. *Nature, 401*(6754), 699–703. <https://doi.org/10.1038/44372>

Toppino, T. C., & Cohen, M. S. (2009). The Testing Effect and the Retention Interval. *Experimental Psychology, 56*(4), 252–257. <https://doi.org/10.1027/1618-3169.56.4.252>

Tsapkini, K., Frangakis, C. E., & Hillis, A. E. (2011). The function of the left anterior temporal pole: evidence from acute stroke and infarct volume. *Brain, 134*(10), 3094–3105.

Tse, C.-S., Balota, D. A., & Roediger III, H. L. (2010). The benefits and costs of repeated testing on the learning of face–name pairs in healthy older adults. *Psychology and aging, 25*(4), 833.

Tse, D., Langston, R. F., Kakeyama, M., Bethus, I., Spooner, P. A., Wood, E. R., Witter, M. P., & Morris, R. G. (2007). Schemas and memory consolidation. *Science, 316*(5821), 76–82.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie canadienne*, 26(1), 1.

Tulving, E. (1995). *Organization of memory: quo vadis?*

Tulving, E., & Markowitsch, H. J. (1998). Episodic and declarative memory: role of the hippocampus. *Hippocampus*, 8(3), 198–204.

Uncapher, M. R., & Wagner, A. D. (2009). Posterior parietal cortex and episodic encoding: insights from fMRI subsequent memory effects and dual-attention theory. *Neurobiology of learning and memory*, 91(2), 139–154.

van den Broek, G. S., Segers, E., Takashima, A., & Verhoeven, L. (2014). Do testing effects change over time? Insights from immediate and delayed retrieval speed. *Memory*, 22(7), 803–812.

van den Broek, G. S., Takashima, A., Segers, E., Fernández, G., & Verhoeven, L. (2013). Neural correlates of testing effects in vocabulary learning. *Neuroimage*, 78, 94–102.

Van Kesteren, M. T., Ruiters, D. J., Fernández, G., & Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends in neurosciences*, 35(4), 211–219.

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.

Verkoeijen, P. P., Tabbers, H. K., & Verhage, M. L. (2011). Comparing the effects of testing and restudying on recollection in recognition memory. *Experimental Psychology*.

Vilberg, K. L., & Rugg, M. D. (2008). Memory retrieval and the parietal cortex: a review of evidence from a dual-process perspective. *Neuropsychologia*, 46(7), 1787–1799.

Visser, M., & Lambon Ralph, M. A. (2011). Differential contributions of bilateral ventral anterior temporal lobe and left anterior superior temporal gyrus to semantic processes. *Journal of Cognitive Neuroscience*, *23*(10), 3121–3131.

Wagner, A. D., Shannon, B. J., Kahn, I., & Buckner, R. L. (2005). Parietal lobe contributions to episodic memory retrieval. *Trends in Cognitive Sciences*, *9*(9), 445–453. <https://doi.org/10.1016/j.tics.2005.07.001>

Walker, M. P., Brakefield, T., Allan Hobson, J., & Stickgold, R. (2003). Dissociable stages of human memory consolidation and reconsolidation. *Nature*, *425*(6958), 616–620. <https://doi.org/10.1038/nature01930>

Weinberger, N. M. (2004). Specific long-term memory traces in primary auditory cortex. *Nature Reviews Neuroscience*, *5*(4), 279–290.

Willoughby, K. A., Desrocher, M., Levine, B., & Rovet, J. F. (2012). Episodic and Semantic Autobiographical Memory and Everyday Memory during Late Childhood and Early Adolescence. *Frontiers in Psychology*, *3*. <https://doi.org/10.3389/fpsyg.2012.00053>

Wills, T. A., Vaccaro, D., & McNamara, G. (1994). Novelty seeking, risk taking, and related constructs as predictors of adolescent substance use: an application of Cloninger's theory. *Journal of substance abuse*, *6*(1), 1–20.

Wing, E. A., Marsh, E. J., & Cabeza, R. (2013). Neural correlates of retrieval-based memory enhancement: an fMRI study of the testing effect. *Neuropsychologia*, *51*(12), 2360–2370.

Wittmann, B. C., Bunzeck, N., Dolan, R. J., & Düzel, E. (2007). Anticipation of novelty recruits reward system and hippocampus while promoting recollection. *Neuroimage*, *38*(1), 194–202.

Wittmann, B. C., Schott, B. H., Guderian, S., Frey, J. U., Heinze, H.-J., & Düzel, E. (2005). Reward-related fMRI activation of dopaminergic midbrain is associated with enhanced hippocampus-dependent long-term memory formation. *Neuron*, *45*(3), 459–467.

Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Annu. Rev. Psychol.*, *55*, 235–269.

World Health Organization (2015). Draft 1: Global strategy and action plan on ageing and health. *Open Consultation: August-October, 2015*, 1–43.

Xiang, J.-Z., & Brown, M. W. (2004). Neuronal responses related to long-term recognition memory processes in prefrontal cortex. *Neuron*, *42*(5), 817–829.

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.

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. <https://doi.org/10.1002/hipo.20864>

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. *Consciousness and Cognition*, *5*(4), 418–441. <https://doi.org/10.1006/ccog.1996.0026>

Yousry, T. A., Schmid, U. D., Jassoy, A. G., Schmidt, D., Eisner, W. E., Reulen, H.-J., Reiser, M. F., & Lissner, J. (1995). Topography of the cortical motor hand area: prospective study with functional MR imaging and direct motor mapping at surgery. *Radiology*, *195*(1), 23–29.

Zald, D. H., Cowan, R. L., Riccardi, P., Baldwin, R. M., Ansari, M. S., Li, R., Shelby, E. S., Smith, C. E., McHugo, M., & Kessler, R. M. (2008). Midbrain dopamine receptor availability is inversely associated with novelty-seeking traits in humans. *Journal of Neuroscience*, *28*(53), 14372–14378.

Zhang, H., Fell, J., & Axmacher, N. (2018). Electrophysiological mechanisms of human memory consolidation. *Nature communications*, *9*(1), 1–11.