

# Recognition memory associated with hippocampal pattern completion in young and older adults

DISSERTATION

zur Erlangung des akademischen Grades

**doctor rerum naturalium**

**(Dr. rer. nat.)**

genehmigt durch die Fakultät für Naturwissenschaften  
der Otto-von-Guericke-Universität Magdeburg

von **M. Sc. Paula E. Vieweg**  
geboren am 16.11.1987 in Bautzen

Gutachter: Prof. Dr. Thomas Wolbers  
Prof. Dr. Jan M. Wiener

eingereicht am: 21.02.2017  
verteidigt am: 20.09.2017

# ACKNOWLEDGEMENTS

First of all, I would like to thank my supervisor Thomas Wolbers, for your guidance and support throughout this process. Thank you for the possibility to work in your lab continuously after the Edinburgh experience, having the opportunity to enjoy the merits of 7T imaging (Oh, so beautiful images!), and despair under its processing nightmares.

Next, my lovely labmates, thank you guys for the ride: Matthias Stangl – my buddy in despair and mutual cheeseboard aficionado; Lorelei Howard – train, Kniffel and office sleepover friend; JonJon Shine – a major bearing on my comic relief; Nadine Diersch – the train rides, parsley pesto and Frau Krause; Martin Fluff-Brain Riemer – the reading, writing and useless discussions.

David Berron, Heldenhafter, I would not have made it without you. My sincerest thanks for your help and willingness to discuss my hypotheses for the hundredth time, but also for the endless other life, science, networking and politics discussions, the conference experiences, the occasional cigarettes, and the beers. Thank you.

I would also like to thank Arturo Cardenas-Blanco for getting me started on coregistration and bash scripting and providing continued support with all of that.

Thank you, Xenia Grande, for discussing interpretations and the literature especially in the final stages, whenever I popped into your office to annoy you or David.

Furthermore, I would like to thank all the people helping with participant recruitment, data acquisition, administration and all the bits providing the groundwork for my research: Henrike Raith, Carla Bilsing, Fana Samatin, Katharina Mamsch, Christin Ruß, Franziska Schulze, Anett Kirmess, Karen Müller-Zabel, Skadi Meister, and Anke Rühling.

A special thanks goes out to the editor of my first paper, Asaf Gilboa, who oversaw a very positive review process and continued to promote my paradigm resulting in a very fruitful collaboration and several new requests for the use of my task. Consequently, I extend more thanks to Steven Baker and Shayna Rosenbaum for the opportunity to contribute to a very intriguing case study.

Being able to test without delay, I want to thank my supportive collaborators in Bonn: Tony Stöcker, Rüdiger Stirnberg and especially Jenny Faber, who put in a lot more effort than originally estimated. Related to this, I would like to thank Myung-Ho In for providing and installing his functional sequence on the scanner in Bonn.

Many thanks to Laura Wisse for intense, very honest and scientifically thorough cooperative work including countless Skype discussions refining our new segmentation rules. This has been a tedious but very rewarding experience. To this end, I would also like to thank Anne Hochkeppler and Anica Luther for their invaluable Sisyphus segmentation effort.

Finally, I want to thank my family, especially my parents Astrid and Klaus, for raising me to be the curious and analytical person who I am. Jakobi and Maresa, thanks for picking me up after the fall. And the biggest thanks to all my friends, you know I mean you, especially Achim, Claudi, Felix, Susi, Hanni, Anna, Fethami, Martin, and Tobi - my man - for bearing with me and reminding me of the really essential bits of life. I am back again now.

# ABSTRACT

Remembering past events is often accomplished with ease, however, when information is scarce this task can become frustratingly difficult. For example, seeing the face of a long absent friend can rapidly trigger a host of related memories to resurface or, if we are unlucky, deficits in this recollective process and the absence of sufficient cues can leave us embarrassed as we draw a blank. Pattern completion is essential for the successful retrieval of such memories. During this process, the original memory trace is restored (completed) via repeated reactivation. Given its extensive excitatory recurrent connections, region CA3 within the hippocampus has been identified as a likely candidate to execute the auto-associative processing essential for pattern completion.

In addition, the structural integrity of the hippocampus is particularly sensitive to the aging process. Specifically, the number of projections from the entorhinal cortex to region CA3 decreases with age, while its interconnectivity remains unchanged. Thus, it is suggested that the aged brain should show a bias toward pattern completion concurrent with CA3-hyperactivity.

In this thesis, I have investigated pattern completion and concurrent changes in aging over the course of four experiments. First of all, I have developed a recognition memory paradigm specifically targeting pattern completion by manipulating stimulus completeness. Simultaneously, age-related recognition memory deficits were identified suggesting a bias towards- but also a deficit in pattern completion. I have further used this paradigm with concurrent eye-tracking which eliminated perceptual contributions to memory performance. From collaborative investigations in a patient with selective bilateral dentate gyrus lesions, first inferences could be derived over the differential contributions of hippocampal subfields to memory performance in the task. Finally, a 7 Tesla ultra-high resolution functional magnetic resonance imaging study yielded controversial findings. While specific hippocampal subfield contributions could not be pinpointed, the whole hippocampus was involved in more general retrieval. Furthermore, the superior temporal sulcus was identified as a region of cortical reinstatement after successful pattern completion. Crucially, age comparisons revealed reduced activity in parahippocampal cortex and hyperactivity in CA3. The latter finding supports existing theories about cognitive aging, and is the first to specifically identify age-related CA3-hyperactivity.

Additionally, to improve existing methodology a segmentation protocol for medial temporal lobe subregions has been developed in a collaborative effort that incorporates recent findings in neuroanatomy.

Altogether, the findings presented in this thesis contribute to the literature on pattern completion and provide a new reliable means of assessing that memory process. The results support and advance existing theories of memory and aging, but also question some of its more specific predictions in MR research. Finally, an important methodological contribution is made to better define medial temporal lobe subregions consistent with the most recent neuroanatomical knowledge.

# TABLE OF CONTENTS

<b>1. General Introduction.....</b>	<b>1</b>
1.1. Hippocampal anatomy.....	1
1.2. Recognition memory.....	2
1.3. Pattern separation and pattern completion.....	3
1.3.1. Computational theories.....	3
1.3.2. Evidence from rodent research.....	4
1.3.3. Evidence from human research.....	5
1.4. Aging.....	8
1.4.1. Aging in the hippocampus.....	8
1.4.2. Aging and recognition memory.....	9
1.4.3. Aging and pattern separation and pattern completion.....	10
1.5. Aim of this thesis.....	12
<b>2. A new task to assess recognition memory associated with pattern completion and the corresponding age-effects.....</b>	<b>13</b>
2.1. Introduction.....	13
2.2. Methods.....	14
2.2.1. Subjects.....	14
2.2.2. Materials.....	14
2.2.3. Procedure.....	15
2.2.4. Bias measure.....	16
2.3. Results.....	16
2.3.1. Accuracy.....	16
2.3.2. Response bias.....	17
2.3.3. Reaction times.....	19
2.3.4. Confidence ratings.....	20
2.4. Discussion.....	20
2.5. Contributions.....	24

<b>3. How do eye-movements contribute to age-related recognition memory differences assessed by the MIC?</b> .....	<b>25</b>
3.1. Introduction .....	25
3.2. Methods.....	26
3.2.1. Subjects .....	26
3.2.2. Materials & procedure.....	26
3.2.3. Eye-tracking acquisition and analysis .....	27
3.3. Results .....	27
3.3.1. Behavioural results.....	27
3.3.2. Eyetracking results.....	31
3.3.3. Summary of the results .....	33
3.4. Discussion .....	33
3.5. Contributions .....	34
<b>4. MIC performance assessed in a patient with bilateral DG lesions</b> .....	<b>35</b>
4.1. Introduction .....	35
4.2. Methods.....	36
4.2.1. Subjects .....	36
4.2.2. Materials & Procedure.....	36
4.3. Results .....	37
4.4. Discussion .....	38
4.5. Contributions .....	39
<b>5. Assessing the neural mechanisms contributing to MIC performance with 7T-fMRI</b> 40	
5.1. Introduction .....	40
5.2. Methods.....	41
5.2.1. Subjects .....	41
5.2.2. Materials.....	42
5.2.3. Procedure .....	42
5.2.4. MRI acquisition.....	43
5.2.5. MRI analysis.....	44
5.3. Results .....	50
5.3.1. Behavioural Results.....	50
5.3.2. Neuroimaging Results .....	52
5.4. Discussion .....	58
5.5. Contributions .....	63

<b>6. Segmentation protocol for MTL subregions in 7T-MRI.....</b>	<b>64</b>
6.1. Introduction .....	64
6.2. Materials and methods .....	66
6.2.1. Participants .....	66
6.2.2. Workshop .....	67
6.2.3. Image acquisition .....	67
6.2.4. Segmentation software .....	67
6.2.5. Manual segmentation protocol .....	67
6.2.6. Statistical analyses.....	84
6.3. Results .....	84
Reliability .....	84
6.3.2. Volumes in comparison to anatomy .....	87
6.4. Discussion .....	88
6.5. Contributions .....	93
<b>7. General discussion.....</b>	<b>94</b>
7.1. Summary.....	94
7.2. Pattern completion targeted by a recognition memory task .....	95
7.3. Recognition memory differences in aging related to pattern completion .....	96
7.4. Implications for general age-related memory performance.....	98
7.5. Brain regions associated with the MIC .....	99
7.6. Outlook and future perspectives.....	101
<b>8. References.....</b>	<b>103</b>
<b>Appendix.....</b>	<b>II</b>
List of abbreviations .....	II
List of figures.....	III
List of tables .....	IV
List of supplementary figures .....	IV
Supplementary figures .....	V
<b>Declaration/Erklärung.....</b>	<b>IX</b>
<b>Paula Vieweg   CURRICULUM VITAE.....</b>	<b>X</b>

# 1. GENERAL INTRODUCTION

The hippocampus has been implicated as a crucial structure for the formation and retrieval of episodic long-term memory – memories of events and their corresponding place and time (Squire et al. 1984; Squire and Zola-Morgan 1991; Squire et al. 2004). The hippocampus is not a homogenous structure but consists of several smaller subfields each implied to fulfil a particular function working in accord to enable the encoding and retrieval of such memories (Marr 1971; Hasselmo and McClelland 1999). Although many memory models exist and empirical knowledge has been accumulating across species over the years, it is still controversial how each hippocampal subfield contributes to these processes (for a recent extensive review, see Kesner and Rolls 2015). Especially in the human, the study of small brain structures is difficult with regard to resolution and the necessity of non-invasive measurements. Crucially, the hippocampus is also very vulnerable to the aging process concomitant with considerable cognitive decline (Small et al. 2011; Leal and Yassa 2015; O’Shea et al. 2016).

In the following paragraphs, I will review hippocampal anatomy along with the most prominent theories about its involvement in memory function with a focus on recognition memory, and pattern separation and completion. Age-related alterations in hippocampal anatomy and function are presented further on.

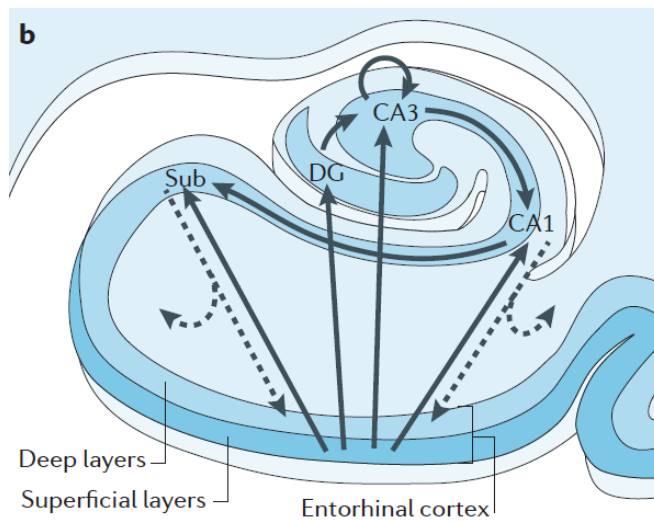
## 1.1. HIPPOCAMPAL ANATOMY

The hippocampus is a neural structure of the limbic system located in the medial temporal lobe (MTL). It consists of several smaller subfields: the cornua ammonis (CA) 1-4 form the hippocampus proper, and constitute the hippocampal formation together with the subiculum and the dentate gyrus (DG) which is composed of the hippocampal hilus and the fascia dentata (Insausti and Amaral 2012). Hereafter, hippocampus refers to the hippocampal formation. Note that the independent existence of CA4 is debated and it is therefore often included into DG, as is here (Amaral 1978; Lorente De N6 1934).

Within the trisynaptic circuit, DG is consecutively connected to CA3 via the mossy fibre pathway, to CA1 via the Schaffer collaterals and to the subiculum (for a recent overview of all connections, see Knierim 2015). CA3 is heavily interconnected through auto-associative fibres, and also projects back to DG (Scharfman 2007). DG neurons portray very sparse firing activity compared to all other hippocampal subfields (Jung and McNaughton 1993).

The hippocampus' main in- and output structure is the entorhinal cortex (ErC; see Figure 1). Its superficial layers project onto the DG and CA3 via the perforant path, providing the hippocampus with sensory input e.g. from perirhinal cortex (PrC), parahippocampal cortex (PhC), auditory and olfactory cortices. The ErC also directly projects to CA1 and subiculum (Steward 1976), but those projections are less well studied. CA1 and subiculum form the main output regions of the hippocampus and

project onto ErC's deep layers (for a comprehensive review, see Derdikman and Knierim 2014).



**Figure 1. Hippocampal connections.** The entorhinal cortex (ErC) projects to DG, CA3, CA1 and subiculum. The major input connection is the perforant path from ErC to DG. From there the trisynaptic pathway connects DG with CA3, CA1 and subiculum. CA1 and subiculum are the major output structures of the hippocampus to the ErC (dashed arrows). The schematic is part of Figure 1 in Small et al. (2011).

## 1.2. RECOGNITION MEMORY

Recognition memory is the ability to retrieve memories of previously encountered events. In one strand of research, it is typically separated into recollection and familiarity (Yonelinas 2002; Diana et al. 2007). The corresponding dual trace model suggests that these processes differently but concurrently lead to recognition and consult different processing streams within the MTL. Familiarity involves unspecific memory of an event that is associated with a sense of vaguely knowing that the event has been previously encountered. Recollection, on the other hand, is thought to entail precise remembering, i.e. a more vivid memory of the event associated with the retrieval of additional information on the context of the event. Therefore, studies typically employ an additional test asking whether participants "know" (familiarity) or "remember" (recollection) a certain event, and to give more information on context or source (where, when or in association with; e.g. Kim and Yassa 2013). While a sense of familiarity is formed rapidly and seemingly effortlessly, the retrieval of more precise or related information associated with recollection is more time-consuming and complex. Findings on the functional localization of these processes are controversial, but more evidence is accumulating on the idea that the hippocampus is involved in recollection, whereas familiarity-based recognition is achieved by the perirhinal cortex (PrC), while other parahippocampal regions seem to contribute to both processes (Aggleton and Brown 2006; Eichenbaum et al. 2007; Yonelinas et al. 2010). However, another account also argues for hippocampal contributions to familiarity (Squire et al. 2004; Wixted and Squire 2010; Smith et al. 2011). It should be noted here, that while familiarity and recollection are mostly discussed as dual processes, there are also other sources advocating a single continuous process where familiarity and recollection are associated with different memory strengths (Slotnick 2013).

Alternatively, the fuzzy-trace theory suggests a slightly different separation into verbatim and gist memory (Brainerd and Reyna 2001). Here, verbatim memory traces



are very detailed and concerned with the "surface form" of the event including context, when a gist trace stores the conceptual information of an event, that is, the attributed meaning or generalization. Crucially, these memory traces are stored simultaneously but verbatim memory is argued to deteriorate faster than gist memory resulting in dissociated availability of each trace during retrieval. False memories are therefore more likely to occur with gist traces, because specific details between exemplars cannot be differentiated from a conceptual memory only. This model was developed to account for developmental changes in false recognition memory (Brainerd and Reyna 2002). Furthermore, the model argues that different types of cues can trigger the two traces distinctly but also complementary; e.g. if the word "collie" is learned, its exact repetition is likely to trigger the verbatim memory trace, however, the more conceptual cue "dog" is more likely to trigger the gist trace, albeit both can influence each other. This leads to another distinction that is relevant in the actual retrieval of memory, that is, recognition vs. recall (Gillund and Shiffrin 1984). While the process of recognition is triggered by a cue that partly or fully resembles the original item, i.e. it is externally aided; free recall has to fully rely on internal information when retrieving a memory. Although distinct, the concepts presented above go hand in hand. Most studies investigating recognition memory employ some sort of cue to trigger memory retrieval, because it can be better controlled. However, the specificity and similarity of the cues vary widely spanning a long range from exact repetitions to greatly distorted items. Consequently, a vague memory matching the concepts of gist memory or the sense of familiarity suffices to trigger recognition of an easy cue. More distorted or partial cues, however, require detailed memory consistent with verbatim memory, in order for correct recollection.

### 1.3. PATTERN SEPARATION AND PATTERN COMPLETION

Memory encoding and retrieval are subject to interference, which occurs when memories associated with the same or a similar cue are competing with each other. To overcome interference, it is necessary that non-identical but similar cues can be associated with a corresponding memory, but similar memories are simultaneously stored as distinct entities. Pattern separation and completion are neural processes suitable to resolve memory interference. Specifically, pattern separation is thought to reduce overlap of representations during memory encoding, while pattern completion restores memory traces from partial or degraded input during memory retrieval. There is a variety of different fields trying to define and investigate these processes; below I review the most prevalent theories.

---

#### 1.3.1. COMPUTATIONAL THEORIES

Derived from hippocampal neuroanatomy, several computational theories have been defined to explain how the interplay of certain regions could achieve rapid and complex memory processing. First described by David Marr (1971) as the "collateral effect", the completion of partial representations could be achieved by co-activation of relevant cells from only a small subset of cells, and suppression of irrelevant cells. This

should best be achieved by an auto-associative network of cells (McClelland et al. 1995; Treves and Rolls 1994). Thus, incomplete or partial memory traces are made more similar (or generalized) to previously stored representations (i.e. pattern completion) and consequently reinstated in the neocortex. There is wide agreement that CA3 is a likely region to perform this sort of processing due to its many excitatory recurrent connections (Treves and Rolls 1994; Marr 1971; McNaughton and Morris 1987; McClelland and Goddard 1996; Hasselmo and McClelland 1999).

Simultaneously, the rapid storage of new representations is suggested to rely on the transfer of input into a dense system where it can be sparsely distributed (McNaughton and Morris 1987; O'Reilly and McClelland 1994; O'Reilly et al. 1998), and additional feedback modification of inputs could enhance separation (Myers and Scharfman 2011) thereby reducing (or orthogonalizing) representational overlap (i.e. pattern separation). Here, DG is the suggested region because of its dense neuron population of approximately 5 times more neurons than projections it receives from ErC neurons, and 6 times more than CA3 neurons it relays to, which as a consequence allows for sparse coding (Amaral and Witter 1989; O'Reilly and McClelland 1994; O'Reilly et al. 1998). Note, that often also the interplay between separation through sparseness in DG and its sparse projections to CA3 is discussed to be responsible for pattern separation (for a detailed review, see Hunsaker and Kesner 2013).

---

### 1.3.2. EVIDENCE FROM RODENT RESEARCH

It has been suggested that memory processes could possibly be better understood by studying hippocampal remapping of place cells (Colgin et al. 2008); cells that fire selectively when the animal is in a particular location (O'Keefe and Dostrovsky 1971; O'Keefe and Nadel 1978). Consequently, most rodent research is done in the spatial domain, i.e. the animals' environment is systematically changed by manipulating local or global landmarks to induce altered hippocampal firing patterns manifesting in either different firing rates of the same cell population (rate remapping) or a global reallocation of place fields (global remapping; Colgin et al. 2008; Fyhn et al. 2007).

Several different hippocampal contributions have been identified in the context of spatial memory. For example, it has been suggested that CA3 can rapidly encode novel sensory input, and integrate the information into an existing framework as a "locally continuous, but globally orthogonal representation" that can then be interpreted by upstream regions like CA1 independent of the spatial context (Leutgeb and Leutgeb 2007). However, CA3 only showed pattern separation after more substantial environmental changes when directly compared to CA1 (Vazdarjanova and Guzowski 2004) or DG (Leutgeb et al. 2007). More specifically, Vazdarjanova & Guzowski (2004) found that when rats were exposed to the same environment twice, CA3 and CA1 ensembles were activated with a similarly high degree of overlap, whereas a completely different environment would produce low overlap, with even less overlap in CA3 compared to CA1. However, changing the identity or configuration of local cues, or changing distal cues activated CA3 and CA1 ensembles with reduced overlap, yet with greater overlap in CA3 than in CA1. Thus, while CA1 exhibited a gradient

response (neuronal overlap decreased with decreasing similarity of the environment), CA3 showed early pattern completion (neuronal overlap was fairly high when the environment was altered but similar to the original) and late increased pattern separation (neuronal overlap was very low when the environments were very different). Similarly, when intra-environmental references were changed rather than the whole environment, population responses in CA3 overlapped more with the original response than CA1 population responses (Lee et al. 2004). The authors suggested that CA1 compares ErC input with separated or completed output from DG and CA3, and thus functions as an integration unit before projecting back to ErC. Further on, gradual environmental changes induced immediate rate remapping of DG neurons with only small changes in CA3 neuron activity possibly counteracted by simultaneous pattern completion, but when the environment was drastically changed completely different cell ensembles were recruited in CA3 (Leutgeb et al. 2007). These findings suggest a pattern separation function for both DG (after small environmental changes) and CA3 (after bigger environmental changes). Lesioning the DG also considerably impaired discrimination of spatial locations lending further support for the notion of DG performing pattern separation (Gilbert et al. 2001). Simultaneously, lesioning CA1 impaired discrimination of the sequence of locations suggestive of temporal pattern separation failure. Another study showed that disabling N-Methyl-D-aspartate (NMDA) receptors in DG influenced CA3 firing rate modulation and led to an impairment in rats to differentiate similar contexts in fear conditioning, also suggesting DG is essential for pattern separation (McHugh et al. 2007). On the other hand, disabling NMDA receptors directly in CA3 did not disrupt spatial encoding and retrieval performance; however, when some of the original cues were removed during retrieval the mutant mice performed worse than controls indicating that CA3 is involved in pattern completion from partial cues (Nakazawa et al. 2002).

Most convincingly, a recent study presented direct evidence for the hypothesis that CA3 performs pattern completion, and DG pattern separation (Neunuebel and Knierim 2014). Simultaneous recordings from both regions during local and global environmental changes showed correlated responses (high overlap) in CA3 but disrupted responses (low overlap) in DG to environmental conflict. This finding has recently been specified in the sense that proximal CA3 (the part closer to DG) showed ensemble dynamics similar to DG, i.e. pattern separation, and the largest part including distal CA3 (the part closer to CA2) and CA2 exhibited the expected pattern completion (Lee et al. 2015).

---

### 1.3.3. EVIDENCE FROM HUMAN RESEARCH

In human research, the memory processes pattern separation and completion can only be approximated given that their definition depends on neuronal computations that cannot be tested non-invasively. For that reason, researchers have tried to develop tasks that accommodate certain underlying assumptions or principles such as interference by manipulating sensory input to induce the respective process (Liu et al. 2016).

Most common and by now frequently replicated is a continuous recognition paradigm, the Mnemonic Similarity Task (MST, previously known as Behavioural Pattern Separation - BPS) originally developed by Kirwan and Stark (2007), which was later changed to an incidental encoding task rather than explicit memory judgement (Bakker et al. 2008). In this task, participants are presented with consecutive images of simple objects that are either new, old (repetitions), or similar to a previous stimulus (lures). For explicit memory testing, participants had to judge the recurrence of each stimulus (new, old, similar), or to parallel implicit tasks from rodent research, make an indoor/outdoor judgement with a post-test recognition memory survey. Critically, concurrent functional magnetic resonance imaging (fMRI) data was analysed taking advantage of the repetition suppression effect which assumes that blood-oxygen-level dependent (BOLD) signal changes when a stimulus is repeated possibly due to adaptation (Krekelberg et al. 2006; for a review, see Grill-Spector et al. 2006). Thus, if pattern completion occurs, i.e. a similar lure is treated as a repetition, activity levels should be similar to ordinary repetitions. In contrast, if pattern separation occurs, i.e. a similar lure is treated as a new stimulus, activity levels should be similar to new stimuli (for a detailed explanation, see Yassa and Stark 2011). The fMRI results revealed CA1 activity indicative of pattern completion, and DG/CA3 activity suggestive of pattern separation (Bakker et al. 2008). These findings were subsequently confirmed with gradual levels of lure similarity (Lacy et al. 2011), as part of comparisons with aged or cognitively impaired populations (Yassa et al. 2011a, 2011b; Bakker et al. 2012), in spatial and temporal versions of the task (Azab et al. 2014), with emotional information (Leal et al. 2014), and in an object/location comparison (Reagh and Yassa 2014a). After the task had been established, it was also used in purely behavioural form, sometimes with slightly different materials but the same general procedure, mainly as a means of comparison between different study populations (Toner et al. 2009; Holden et al. 2013; Ally et al. 2013; Stark et al. 2013; Baker et al. 2016) or task demands (Duncan et al. 2012; Motley and Kirwan 2012; Kim and Yassa 2013; Stark et al. 2015).

Since then, a few other tasks have been developed to tackle pattern separation and completion in humans. For example, Bonnici et al. (2012a, 2012b) have devised an fMRI paradigm with morphed stimuli, i.e. two distinct but highly similar mountain ranges were created with seven additional gradual morphs between the two (e.g. 70% of range A and 30% of range B, 50% of range A and 50% of range B). After learning, participants had to decide for each stimulus whether it was more likely A or B, and were rewarded or punished monetarily depending on their performance (albeit feedback being only given during training). The authors employed multivariate pattern analysis (MVPA) to classify brain activity corresponding to the two mountain ranges and reported that the hippocampus could decode the identity of the 100% stimuli, which was taken to reflect appropriate pattern separation (Bonnici et al. 2012b). More specifically, all hippocampal subfields could differentiate between the participants' decisions for 100% stimuli, but CA1 and CA3 stood out in classifying the decisions on the 50% morphed stimuli (Bonnici et al. 2012a). The classification for decisions was reasoned to depend on pattern completion, that is, when the stimulus equally consisted

of both mountain ranges (50% morphs), accurate classification necessarily depended on an additional process leading to the decision for one or the other mountain range. It should be noted that seemingly all subfields could do this with very high classification accuracies (in MTL standards) of over 70%, but CA1 and CA3 were comparably the highest with over 80%. This was the first study to suggest that differences between hippocampal subfields, and specifically between CA3 and DG, could be detected with MVPA classification (Bonnici et al. 2012a).

Another strand of research tried to find human analogues for typically used paradigms in rats by using a virtual environment (Paleja et al. 2011). After participants were excessively trained to navigate inside a square environment, they had to find a specific location amongst several differently spaced options to trigger pattern separation, or with less distal cues available than during training to induce pattern completion. Targets that were closer together in the separation task were less often found than if they were further apart suggesting that pattern separation demands increased. Similarly, performance decreased when less environmental cues were available suggesting increased pattern completion demands. Their study may provide a useful translation from animal to human research.

Meanwhile, another paradigm has been developed to assess pattern completion in the human (Horner and Burgess 2014), which was not available at the time this dissertation project was developed. It assumes that memories are stored in so-called engrams – coherent representations of multiple elements of an event such as location, person and object. In their experiments, unique combinations of these three categories had to be learned either all together ("simultaneous" condition) or in subsequent pairs, linking all three elements with each other element ("separated closed loop" condition), or linking the three elements with a fourth element but not using all possible pairs ("separated open loop" condition), e.g. location-object, person-animal, location-person, but not object-animal. Memory was then tested with a cued-recognition task where one element was presented and a corresponding element had to be picked out of six options within the same category. Although participants performed slightly worse in the open-loop condition compared to all other conditions, analyses mainly dwelled on comparisons of within-condition dependencies with dependent and independent models. Dependencies reflected "the proportion of events in which both associations were retrieved correctly or both incorrectly" (Horner and Burgess 2014). Simultaneous and closed-loop condition dependencies were comparable to the dependent model, while open-loop condition dependencies were similar to the independent model. The authors argue that the closed-loop condition allowed for pattern completion, as all elements within one engram were similarly well remembered, i.e. showed dependency, while this was not true for the open-loop condition. There, retrieval was better for pairs that had been studied together, thus no coherent representation of the full engram had been established, therefore pattern completion was absent. In summary, the authors suggested that associations between all elements of an engram allow retrieval via pattern completion.

## 1.4. AGING

Aging is accompanied by a number of structural and behavioural changes; e.g. knowledge increases, bodily functions deteriorate, and general processing speed decreases (for review, see Park and Festini 2016). One prominent consequence of aging is memory decline associated with structural brain changes in the MTL, and the hippocampus in particular. Here, I review alterations in the hippocampus, along with changes in recognition memory, and pattern separation and completion more specifically.

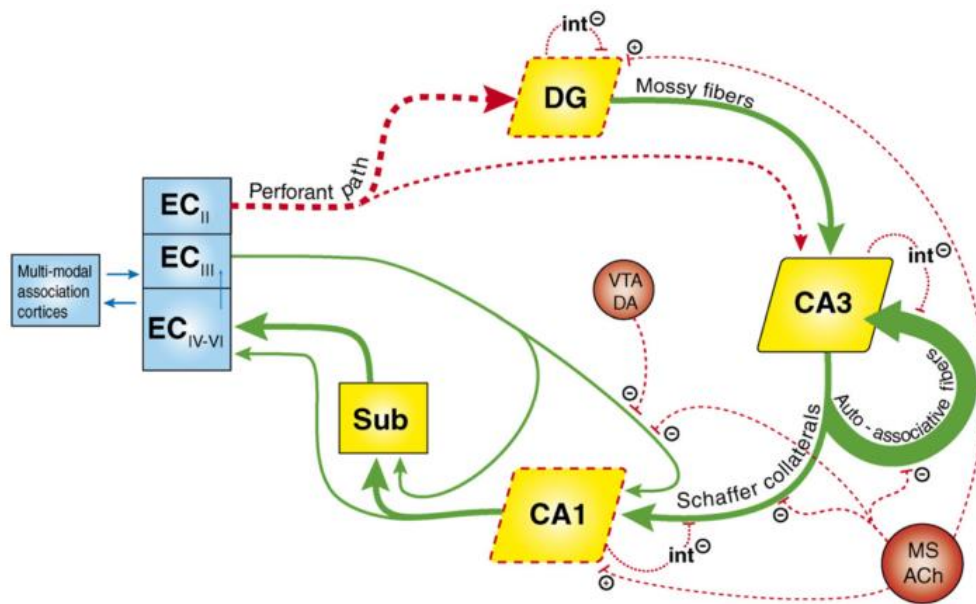
---

### 1.4.1. AGING IN THE HIPPOCAMPUS

With age, the hippocampus undergoes severe structural changes; its volume decreases substantially (Raz et al. 2005; for a recent meta-analysis, see Fraser et al. 2015), these alterations are subfield-specific (de Flores et al. 2015b; Wilson et al. 2005), and they affect cognition (O'Shea et al. 2016), and especially memory (Yassa et al. 2011b; Shing et al. 2011; Travis et al. 2014; for a recent exemplary longitudinal study, see Gorbach et al. 2016).

The main underlying circuit changes as investigated in animals include (1) a degradation of the perforant path (Smith et al. 2000), thus, a reduction of sensory input from ErC into the hippocampus, (2) reduced cholinergic modulation in CA3 and CA1 (Hasselmo et al. 1995; Nicolle et al. 2001), (3) reduced dopaminergic modulation in CA1 (Hemby et al. 2003), (4) reduced inhibitory interneuron activity in CA1 (Vela et al. 2003; Stanley and Shetty 2004), and (5) overall weakened synaptic plasticity (Burke and Barnes 2006). These factors have been argued to contribute to reduced excitability in DG and CA1 (Burke and Barnes 2006), hyperactivity and rigidity in CA3 cells (Wilson et al. 2005), and additionally, the DG is pinpointed as the most vulnerable subfield in healthy aging compared to pathological aging and disease (Small et al. 2004, for a comprehensive review, see 2011). However, importantly, it should be noted that in spite of the manifold findings from the animal literature, research on healthy aging in humans is controversial regarding the specific contributions and volume changes in hippocampal subfields, i.e. the identification of different subfields and accompanying cognitive changes varies considerably between different studies (for a comprehensive review of controversial findings, see de Flores et al. 2015b).

Mostly based on the disintegration of the perforant path, and the CA3 auto-associative connections remaining intact (Smith et al. 2000), one aging model suggests that circuit-specific disruptions in the hippocampus lead to an alteration in the memory system by favouring the retrieval of previously stored events and putting the encoding of new events to a disadvantage (Wilson et al. 2006). In detail (for a descriptive schematic, see Figure 2), as the perforant path degenerates, DG and CA3 receive less sensory input. This leads to hypoactivity in DG, and reduced cholinergic input to CA3. At the same time, interconnections within CA3 remain unimpaired. Following these premises, CA3 reactivates its auto-associative cells based on less sensory input, and on less pre-processed output from DG, thereby having to rely more on internal information. This results in a hippocampal network preference of retrieval over encoding.



**Figure 2. Hippocampal degeneration in aging.** This schematic is taken from Figure 1 in Wilson et al. (2006). Age-related alterations of the hippocampal circuit are shown with red dashed lines, connections and regions that remain intact, or where there are not sufficient data are depicted in green. Crucially, (1) the perforant path degenerates, (2) there is reduced cholinergic (ACh) modulation by the medial septum (MS) and (3) reduced dopaminergic (DA) modulation by the ventral tegmental area (VTA), (4) inhibition by interneurons (int) is decreased, and (5) excitability in DG and CA1 is reduced. All these changes may contribute to hyperactivity in CA3 auto-associative connections (thick green arrow).

#### 1.4.2. AGING AND RECOGNITION MEMORY

The structural changes of the aging brain and the hippocampus in particular are accompanied by cognitive decline especially in the memory domain. While semantic memory seems to be preserved with age (Park et al. 2002), episodic memory, working memory, processing speed, and executive function are impaired (Leal and Yassa 2015; Jagust 2013; Grady 2012; O'Shea et al. 2016). However, there are considerable differences in decline regarding the type of episodic memory. That is, there is a variety of findings which indicate that recollection is primarily impaired with age and familiarity is mostly spared (for a recent meta analysis, see Koen and Yonelinas 2014), however, there are contradicting results indicating that familiarity can also be affected (e.g. Duarte et al. 2010). As an example, older adults are able to indicate that they have seen an item when primed with it and having to make a yes/no judgement which suggests that familiarity is intact. However, when asked to relate the item to a source, they are impaired in doing so indicative of a deficit in recollection (Koen and Yonelinas 2016). Similarly, older adults more often choose an unspecific "know" option when being asked to indicate whether they remember or know an item or if it is new when compared to a younger group (Duarte et al. 2010). Next to inadequate source judgements, these findings receive further support from the literature discussing different ways of retrieval. More specifically, free recall is considerably more impaired in older adults than when retrieval is aided by a particular cue (for review, see Lindenberger and Mayr 2014). For illustration, when participants have learned a list of words and are later asked to freely recall any word they still remember, older adults remember notably fewer words than young adults. However, when primed with the word and having to indicate if they have learned it or not, their impairment is less

prominent (Danckert and Craik 2013). These results go a long way in suggesting a general age-related deficit in self-initiated processing as free recall requires and the concomitant need to rely more on environmental information (Luo and Craik 2008; Lindenberger and Mayr 2014).

---

### 1.4.3. AGING AND PATTERN SEPARATION AND PATTERN COMPLETION

The model of neurocognitive aging by Wilson et al. (2006; as presented above) suggests that the aged memory system should show a bias toward pattern completion, and a concurrent deficit in pattern separation. This conceptualization has been taken up by many scientists trying to investigate these processes in human aging. The first study trying to address these questions, employed the MST with explicit memory judgements (as in Kirwan and Stark 2007; see task description above 1.3.3) in healthy young and older adults (Toner et al. 2009). While older adults performed just as well as young adults in the identification of old and new items, they were impaired in the correct identification of similar items (lures) and tended to think they were old instead (false alarms). This was taken to suggest that pattern separation was less efficient in aging. Consequently, the next studies tried to identify the underlying neural changes to this behavioural impairment. Using the same task again, participants were simultaneously scanned with fMRI by which the CA3/DG complex was observed to exhibit hyperactivity that correlated with the specific performance deficit in older adults (Yassa et al. 2011a). Simultaneously, performance on different levels of lure similarity was assessed, showing that older adults always performed worse than young adults, that is, they needed a bigger perceptual change in the stimulus to recognize it was only similar and not identical to an old one. This was interpreted to be in line with the model presented above (Wilson et al. 2006) in that the system was biased towards pattern completion rather than pattern separation. In further support of this model, the gradual change over different levels of lure similarity could also be followed by a corresponding activity profile in the CA3/DG complex coined "representational rigidity" which correlated with structural changes of the DG/CA3 complex and the perforant path assessed by diffusion imaging (Yassa et al. 2011b; see also Bennett and Stark 2016). These studies were the first to employ one value, the now termed Lure Discrimination Index (LDI), to assess performance biases, which is now reported for results of these paradigms as a standard feature. It is calculated as the difference between similar responses to lure items and similar responses to new items accounting for a general response bias; higher values are taken to indicate better pattern separation efficiency. Follow-up studies revealed that this measure was more sensitive to cognitive decline than standard recognition memory (Holden et al. 2013; Stark et al. 2013). More specifically, older adults were grouped into impaired and unimpaired individuals based on a delayed verbal recall task. Crucially, the unimpaired older group performed just as young adults whereas the impaired group had lower LDI scores (Holden et al. 2013). Moreover, the scores of the impaired group were similar to that of MCI patients, while this difference did not show in a standard recognition memory comparison (Stark et al. 2013). Equally, temporal order judgments of consecutively presented objects revealed



that older adults performed generally weaker than young adults, but only the impaired group showed performance suggestive of a pattern separation deficit (Roberts et al. 2014). In a verbal version of the task, older adults were impaired in phonological similarity discrimination (if words sounded/looked similar), but not in semantic discrimination (if words had similar meanings), which was interpreted as a pattern separation failure in perceptual discrimination (Ly et al. 2013). Similar findings were reported by two distinct studies transferring this task into the spatial domain; impaired older adults performed worse than young and unimpaired older adults in judging the location of an object as compared to a previous presentation especially when the target was close (more similar) to its original presentation (Holden et al. 2012; Reagh et al. 2013). Finally, a study using a virtual environment showed that older adults were worse at navigating to a target location when fewer extra-maze cues were present pointing towards an age-related deficit in pattern completion (Paleja and Spaniol 2013). Albeit being useful in the translation from animal to human research, the design left open how participants were to solve the task when no extra-maze cues were available because the environment did not include any other landmarks.

## 1.5. AIM OF THIS THESIS

Behavioural paradigms have been developed to approximate the neural computations of pattern separation and completion while simultaneously disentangling the involvement of different hippocampal subfields. Given their behavioural nature, however, paradigms are embedded in more psychological concepts like recognition memory or mnemonic discrimination. Thus, although there is an increasing body of literature advancing our knowledge of memory processing in humans, several limitations remain. First of all, there is mainly one paradigm to investigate pattern separation and completion in humans - the mnemonic similarity task (MST; Stark et al. 2015). Albeit successful replication, this task is specifically designed to target pattern separation, i.e. by increasing stimulus similarity, the need to orthogonalize the corresponding representations during encoding is boosted. Pattern completion on the other hand, is merely a by-product of this assessment, observed in failures to dissociate similar stimuli. However, by definition, pattern completion should be accountable for more than behavioural errors, i.e. its important role in retrieval needs to be further illuminated. Note, that since the beginning of this dissertation work a few other tasks have been developed which were not available during the initial planning of the project (Staresina et al. 2013a; Horner and Burgess 2014).

Second, due to neuroimaging constraints with regard to resolution, the distinction of hippocampal subfields has been limited, specifically in the differentiation of CA3 and DG. Most of the human literature has considered a CA3/DG complex with the exception of Bonnici et al. (2012a, 2012b), even though anatomical findings and functional models suggest clearly distinct if not opposing functions for the two regions (McClelland et al. 1995).

Third, the role of aging has received increasing attention over the past years, but alterations in hippocampal anatomy are still controversial (de Flores et al. 2015b). Additionally, the interplay of related functions again hinges on more fine-grained differentiation limited by imaging resolution (Wilson et al. 2006).

This thesis aims to address these limitations over the course of several experiments as follows:

- Chapter 2:** Development of a behavioural task targeting pattern completion and assessing age-related performance differences
- Chapter 3:** Replication of the developed task and elimination of perceptual confounds in memory processing with regard to age differences
- Chapter 4:** Identifying the contribution of hippocampal subfields to task performance: insights from a patient with bilateral DG lesions
- Chapter 5:** Investigation of the neural mechanisms involved in solving the task and age-related performance differences using 7T-fMRI
- Chapter 6:** Development of an improved segmentation protocol consistent with neuroanatomical region differentiation in the MTL

## 2. MEMORY IMAGE COMPLETION

### A new task to assess recognition memory associated with pattern completion and the corresponding age-effects

The experiment presented in this chapter has been previously published as is, only figures, tables and corresponding captions have been edited, and the task has later been named Memory Image Completion (MIC):

Vieweg P, Stangl M, Howard LR, Wolbers T (2015). Changes in pattern completion - A key mechanism to explain age-related recognition memory deficits? *Cortex* 64: 343–351.

#### 2.1. INTRODUCTION

All too often we find ourselves faced with the problem of recognizing something familiar even though its appearance may have changed; for example, finding our way across a park with all the trees having lost their leaves, or recognizing a person wearing a different haircut. Pattern completion is essential for the successful retrieval of memories from such degraded or partial cues. This process has been defined as a hippocampal computation during which the original memory trace is restored (completed) via reactivation (Marr 1971; McClelland et al. 1995). However, behavioural evidence for such computations in episodic memory processing in humans is rare. One line of evidence comes from studies using continuous object recognition tasks to assess pattern separation – a concurrent process which differentiates new input from stored representations (for review, see Yassa and Stark 2011). Typically, stimuli used in these paradigms are similar lures, and participants' ability to correctly reject them as similar and not identify them as old is interpreted as behavioural pattern separation (Stark et al. 2013). The identification of pattern completion processes is usually a by-product of this assessment; that is, the failure to correctly reject a lure as similar and judging it as old (false alarms) is interpreted as behavioural pattern completion (Ally et al. 2013). However, as of yet, it is unclear how exactly pattern separation and completion contribute to behaviour, and whether they are distinct processes that work concurrently or in competition, or whether they represent two ends of a unified process (for review, see Hunsaker and Kesner 2013).

Because the structural integrity of the hippocampus is particularly sensitive to the aging process, it has been suggested that the aged brain should show a bias toward pattern completion (Wilson et al. 2006). Behaviour concomitant with these age-related changes in hippocampal processing has been assessed with a similar focus on pattern separation, only indirectly showing a shift towards pattern completion (Toner et al. 2009; Yassa et al. 2011b). However, a more recent study has raised objections to these conceptualizations by showing that both measures (lure correct rejections and false

alarms) likely entail both pattern separation and completion, suggesting that more process-pure behavioural measurements need to be developed (Molitor et al. 2014). In that study, eye-tracking data revealed that performance differences were driven by differential encoding rather than retrieval, hence lure correct rejections and false alarms should rather be interpreted as successful and unsuccessful pattern separation during encoding as opposed to pattern completion biases during retrieval.

In the present study, we devised a behavioural paradigm more suitable to assess pattern completion, and to test the hypothesis that older adults would show a bias towards this process. We developed a recognition task that required participants to learn simple line-drawn scenes and later identify them amongst new scenes. During recognition, we manipulated stimulus completeness by gradually reducing scene information similar to Gollin figures (Gollin 1960). The resulting partial input was intended to trigger the pattern completion process, a manipulation suggested by Hunsaker and Kesner (2013). With this paradigm, we could (1) assess the recognition ability across different levels of stimulus completeness, and (2) calculate a response bias score by comparing the performance for learned versus new stimuli, while simultaneously characterizing age effects.

## 2.2. METHODS

### 2.2.1. SUBJECTS

All participants were recruited by the German Center for Neurodegenerative Diseases (DZNE) in Magdeburg. After screening for mild cognitive impairment (MCI) using the Montreal Cognitive Assessment (MoCA; Nasreddine et al. 2005), 4 older participants were excluded, because they scored lower than 23 (Luis et al. 2009). Thirty young (20-35 years old; 15 males) and 30 older adults (62-78 years old; 15 males) were included in the study. Informed consent was obtained in writing before the experiment, and the study received approval from the Ethics Committee of the University of Magdeburg. All participants received monetary compensation of 6.50€/h.

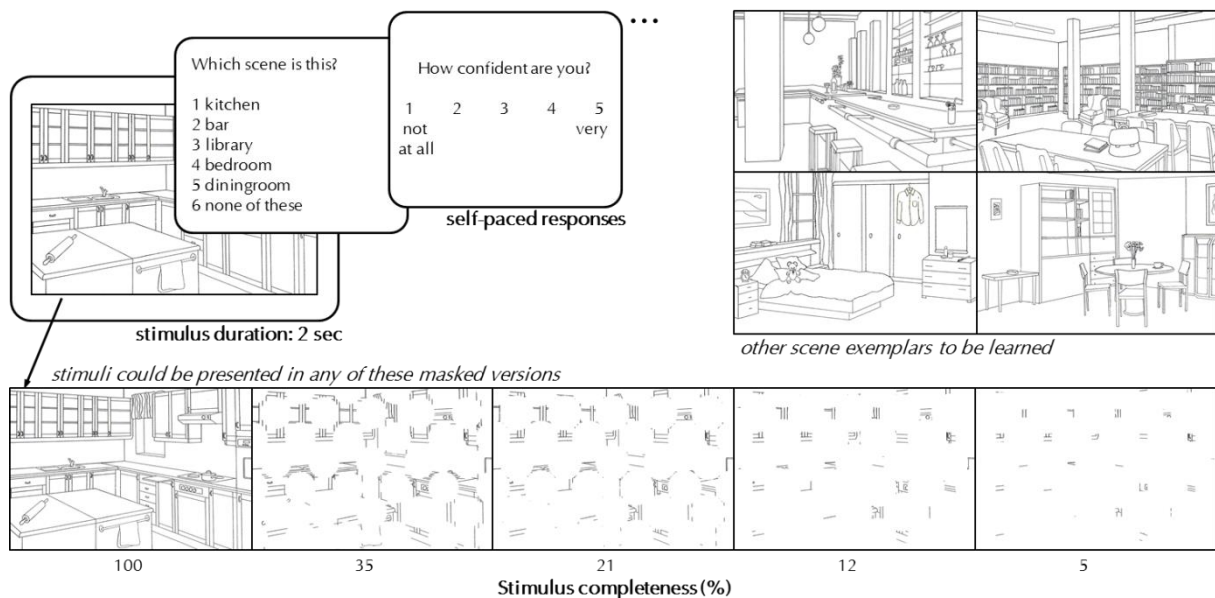
### 2.2.2. MATERIALS

The experimental stimuli comprised 15 black and white line-drawn images (Hollingworth and Henderson 1998) depicting simple indoor scenes (e.g. kitchen, bar, library, etc.). Stimulus completeness was manipulated for 10 of the 15 line-drawn images by masking them with a grid (5x6) of white circles. Four different completeness levels (35%, 21%, 12%, and 5%; percentages reflect the amount of the image visible through the mask) were created by gradually increasing the circle by a factor of 1.2 after each iteration (the size of this manipulation was determined by careful piloting of the paradigm). The original stimulus (100%), therefore, became progressively more occluded by the mask and appeared less complete (see Figure 3, bottom panel). All stimuli were presented on a 15-inch computer screen.

### 2.2.3. PROCEDURE

Prior to the test phase of the experiment (the results of which are outlined in this paper), participants learned 5 different scene exemplars. Each exemplar was presented for 2 seconds in the centre of the screen, on a grey background; a verbal label of the image (e.g. 'dining room') preceded each scene for 1 second. All items were presented 3 times in a random order throughout the learning phase. To ensure that participants remembered the 5 scene exemplars, these items were presented again, intermixed with 5 new scene foils. Each stimulus was presented for 2 seconds, after which participants were required to indicate whether they had seen it before; if so, they had to select the corresponding description from among 3 semantically similar options (e.g. 'kitchen', 'canteen', 'cafeteria'). Participants were allowed to proceed with the experiment only after correctly identifying each learned scene on 3 consecutive trials.

In the test phase of the experiment (see Figure 3), the 5 original scene exemplars were again presented intermixed with 5 novel scene items; all stimuli were presented unmasked (100%) and in the 4 incomplete versions (35%, 21%, 12%, and 5%), resulting in 50 test items. Each item was shown 4 times in a random order with a duration of 2 seconds. On each trial, participants had to indicate which of the 5 learned scenes was presented or whether it was a new scene (i.e. 'bar', 'library', 'dining room', 'bedroom', 'kitchen', 'none of these'). Responses were self-paced. Performance was scored as correct only when participants identified the one appropriate response (i.e. the exact stimulus name for learned stimuli, and 'none of these' for new stimuli), resulting in a chance level of 1/6 for each trial. Additionally, participants had to rate their confidence in this decision on a scale from 1 ('not at all confident') to 5 ('very confident').



**Figure 3. Ch. 2 - design of the test phase.** Each stimulus was presented for 2 s each, followed by 2 self-paced forced choice tasks - stimulus identification and confidence rating. In a previous study phase, participants learned the 5 depicted stimuli (kitchen, bar, library, bedroom, diningroom; from Hollingworth & Henderson, 1998). Those were then mixed with 5 novel items and all 10 were randomly presented in complete or masked form as shown in the bottom panel; percentages reflect the amount of the image visible through the mask.

---

#### 2.2.4. BIAS MEASURE

Test performance for both learned and new items can rely upon pattern completion. The identification of learned items from partial cues (35%, 21%, 12%, and 5%) provides a demonstration of this process. Similarly, to identify new items, participants might employ a recall-to-reject strategy, whereby they retrieve a learned stimulus to compare it to the current sensory input before deciding on whether it is in fact new or not; this strategy, therefore, also relies upon pattern completion. It should be noted, that at the same time pattern separation is likely required to compare and orthogonalize the new item to the retrieved one.

First, performance scores for the learned stimuli (i.e. correctly selecting the exact stimulus name as a response) were obtained, which served as an index of the individual recognition ability. It was then assessed by how much performance for the new stimuli (i.e. correctly selecting 'none of these' as a response) deviated from this value, to test whether there was behavioural evidence for a response bias in older adults. Therefore, the difference in accuracy scores for learned minus new stimuli was calculated separately for each participant and for each level of stimulus completeness. Positive-going values were obtained if a participant's performance for new stimuli was worse than for learned stimuli. This pattern of performance is indicative of a higher tendency to select one of the five learned options when presented with a new stimulus, which is interpreted here as a bias to complete towards a familiar pattern.

---

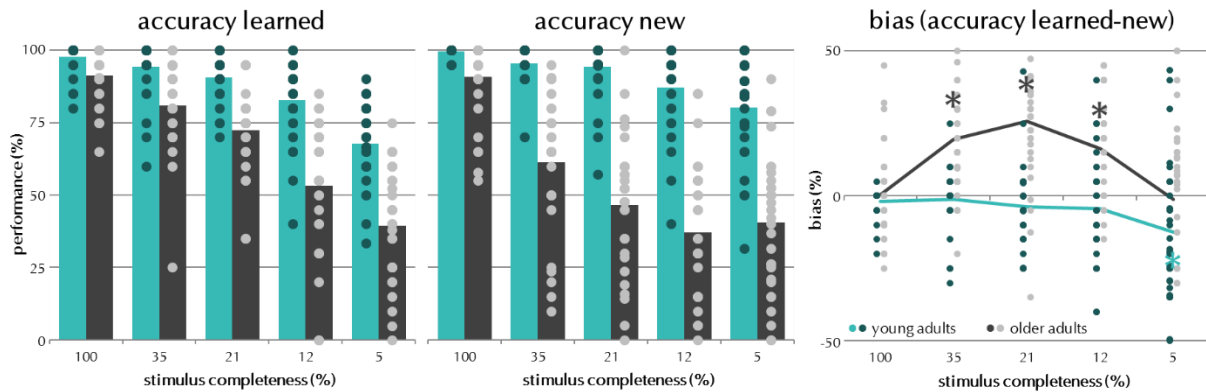
### 2.3. RESULTS

---

#### 2.3.1. ACCURACY

First, recognition ability was assessed by computing accuracy scores separately for learned and new stimuli (see Figure 4, left panel). A three-way mixed analysis of variance (ANOVA) with a between-subjects factor of age (young, old), and two within-subjects factors (stimulus completeness: 100%, 35%, 21%, 12%, 5%; stimulus type: learned, new) revealed that young participants performed better than older participants (main effect of age:  $F_{(1,58)} = 128.342$ ,  $p < 0.001$ ). For both groups, performance was modulated by the degree of stimulus completeness (main effect of stimulus completeness:  $F_{(4,232)} = 256.981$ ,  $p < 0.001$ ), i.e. reduced stimulus completeness resulted in less accurate performance. This decrease was more pronounced in the elderly, as was revealed by a two-way interaction (age  $\times$  stimulus completeness:  $F_{(4,232)} = 46.104$ ,  $p < 0.001$ ). Interestingly, performance was differentially affected between age groups relative to whether they saw a learned or a new stimulus (age  $\times$  stimulus type  $\times$  stimulus completeness:  $F_{(4,232)} = 5.54$ ,  $p < 0.001$ ). In fact, even though the performance per stimulus type (learned, new) was not different overall (main effect of stimulus type:  $F_{(1,58)} = 3.517$ ,  $p = 0.066$ ), the two-way interaction of age and stimulus type showed that only older adults performed worse for new stimuli as compared to learned stimuli (see Figure 4; age  $\times$  stimulus type:  $F_{(1,58)} = 18.227$ ,  $p < 0.001$ ). Post-hoc independent t-tests revealed age group differences in performance for all levels of stimulus completeness for both learned and new stimuli (after Holm-Bonferroni multiple

comparisons correction; all  $p < 0.001$ ; level 100%:  $t_{learned(58)} = 3.397$ ,  $t_{new(58)} = 3.416$ , level 35%:  $t_{learned(50.248)} = 4.059$ ,  $t_{new(58)} = 7.695$ , level 21%:  $t_{learned(45.464)} = 6.125$ ,  $t_{new(58)} = 10.348$ , level 12%:  $t_{learned(58)} = 6.227$ ,  $t_{new(58)} = 10.195$ , level 5%:  $t_{learned(58)} = 6.45$ ,  $t_{new(58)} = 7.38$ ). Altogether, these findings show that older adults' recognition ability was impaired across all levels of stimulus completeness in relation to young adults, and even more so for new stimuli as compared to learned ones.



**Figure 4. Ch. 2 - performance and bias measures.** Left, performance for both age groups, separately for learned and new stimuli for the 5 different levels of stimulus completeness (mean); right, bias measure (see Methods for a detailed explanation) - difference in accuracy scores for learned minus new stimuli calculated separately for each participant (mean); positive values indicate a bias toward pattern completion, significant differences from 0 are indicated with \* separately for each age group as indicated by colour.

### 2.3.2. RESPONSE BIAS

Because I was interested in the identification of response biases, I looked at the distribution of response errors for learned items only. If false familiar responses (false alarms) occur more often than false 'new' responses (misses), this could potentially reveal a pattern completion bias. Numerically, older adults had higher false alarm rates than misses, while the reverse was true for young participants (descriptive statistics can be viewed in Table 1). However, the proportion of errors was too small for a detailed analysis.

**Table 1. Ch. 2 - false alarm rates for learned stimuli.**

stimulus completeness	false alarms - mean (SE)	
	young adults	older adults
100%	0.22 (0.16)	0.35 (0.09)
35%	0.24 (0.10)	0.67 (0.07)
21%	0.39 (0.65)	0.65 (0.05)
12%	0.63 (0.07)	0.53 (0.04)
5%	0.44 (0.05)	0.54 (0.05)

False alarms and misses add up to 1, so that values higher than 0.5 indicate more false alarms, and values lower than 0.5 indicate more misses; values do not comprise the data of all participants since not all of them made errors for each completeness level.

To investigate a potential bias in more detail, accuracy scores for learned stimuli were treated as indices of the individual recognition ability, and I then assessed how much the performance for new stimuli deviated from this index. Therefore, I calculated individual bias scores by subtracting the accuracy scores for new stimuli from the learned stimuli for each participant separately (see chapter 2.2.4 for details). The

resulting bias measures were submitted to a mixed ANOVA (age  $\times$  stimulus completeness). Older adults had higher scores than young adults, i.e. a positive bias (see Figure 4, right; main effect of age:  $F_{(1,58)} = 18.227, p < 0.001$ ). The bias scores were influenced by stimulus completeness (main effect of stimulus completeness:  $F_{(4,232)} = 11.09, p < 0.001$ ), and the two-way interaction with age was also significant (age  $\times$  stimulus completeness:  $F_{(4,232)} = 5.54, p < 0.001$ ). To explore this interaction in more detail, I performed five planned post-hoc comparisons. Independent t-tests demonstrated significant between-group differences only for the middle three completeness levels after Holm-Bonferroni multiple comparisons correction (level 35%:  $t_{(58)} = -4.82, p < 0.001$ , level 21%:  $t_{(58)} = -6.216, p < 0.001$ ; level 12%:  $t_{(58)} = -3.016, p = 0.004$ ).

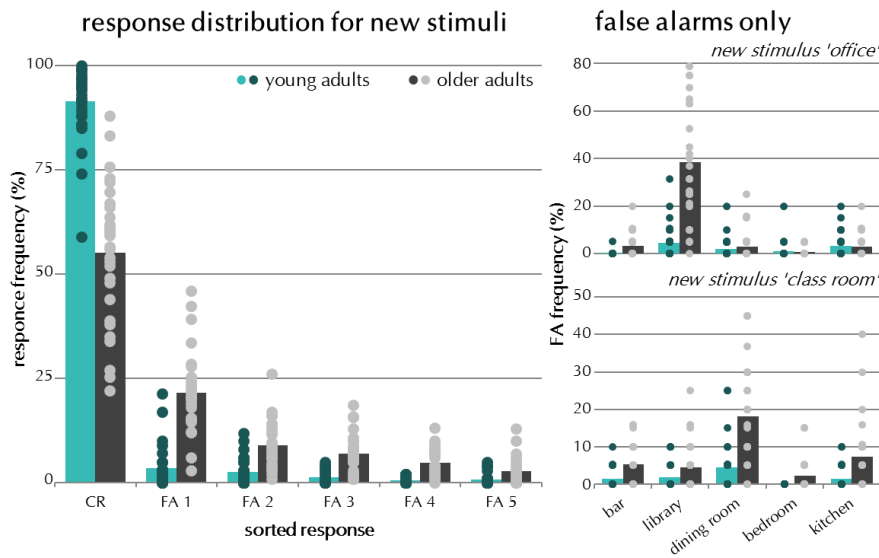
To test the levels of stimulus completeness at which this score establishes a bias, group average bias scores for both the young and older adults were tested against 0 with five one-sample t-tests. Only the older adults showed a positive bias for the middle three completeness levels, indicative of a pattern completion bias (after Holm-Bonferroni multiple comparisons corrections; level 35%:  $t_{(29)} = 5.131, p < 0.001$ , level 21%:  $t_{(29)} = 6.466, p < 0.001$ ; level 12%:  $t_{(29)} = 2.717, p = 0.011$ ), while younger participants showed the opposite, negative bias with the least complete stimuli (level 5%:  $t_{(29)} = -2.868, p = 0.008$ ). There was no evidence of a bias for the complete versions of the stimuli (level 100%:  $t_{young(29)} = -2.009, p_{young} = 0.054$ ;  $t_{old(29)} = -0.145, p_{old} = 0.886$ ). This was to be expected as participants were allowed to continue to this part of the experiment only if they had demonstrated accurate memory for the stimuli and should therefore be able to discriminate learned from new stimuli equally well. An alternative explanation for the lower performance of older adults for new stimuli could be that they guess more. Therefore, I looked at the distribution of errors for each specific new stimulus. If participants were simply guessing, then each false response choice (i.e. the learned stimuli's labels 'bar', 'library', 'dining room', 'bedroom', 'kitchen') should occur equally frequent.

Figure 5 (right panel) shows the false alarm distribution for 2 exemplary stimuli. In fact, especially older adults chose one answer a lot more often than any of the other responses (i.e. 'library' was the most frequent false choice for stimulus 'office', and 'dining room' was the most frequent false choice for stimulus 'class room'). This indicates that the participants were not randomly guessing. Instead, they chose the 1 of the 5 learned stimuli that was presumably perceived as most similar to the new one. To show that this error pattern was not simply driven by those 2 examples, the group-average frequencies of all false choice options was used and sorted from most (FA 1) to least (FA 5) chosen option for each new stimulus. Subsequently, I averaged across all new stimuli and obtained a generalized response distribution differentiating false alarms. The left panel of

Figure 5 shows that older adults indeed chose one particular false response option most often (FA 1) and did not simply guess more, which was confirmed by a  $\chi^2$ -test of goodness-of-fit on the 5 false alarm options ( $\chi^2 = 716.949, df = 4, p < 0.001$ ). Additionally, for each specific new stimulus, the most frequent false alarm was tested



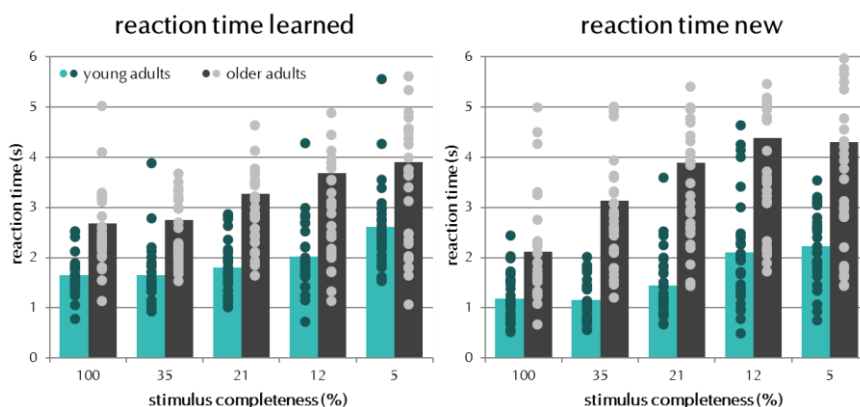
against the average of the other false alarms to show that there was one dominant option per stimulus (stimulus 'office':  $\chi^2 = 185.719$ ,  $df = 1$ ,  $p < 0.001$ ; stimulus 'class room':  $\chi^2 = 44.899$ ,  $df = 1$ ,  $p < 0.001$ ; stimulus 'restaurant':  $\chi^2 = 44.024$ ,  $df = 1$ ,  $p < 0.001$ ; stimulus 'locker room':  $\chi^2 = 15.791$ ,  $df = 1$ ,  $p < 0.001$ ; stimulus 'living room':  $\chi^2 = 9.391$ ,  $df = 1$ ,  $p = 0.002$ ). This indicates that older adults completed towards the stimulus perceived as most similar.



**Figure 5. Ch. 2 - response distribution for new stimuli.** Left, responses are depicted over the 6 possible choice options (i.e., 'none of these' as correct rejections - CR, and the false alarms sorted according to frequency - FA 1-5; mean) showing that older adults chose one particular false response option most often (FA 1) rather than guess more overall, which would lead to similar frequencies for all 5 response options. Right, distributions of false alarms are depicted for 2 exemplary stimuli per actual false response option (i.e., label of the learned stimuli; mean).

### 2.3.3. REACTION TIMES

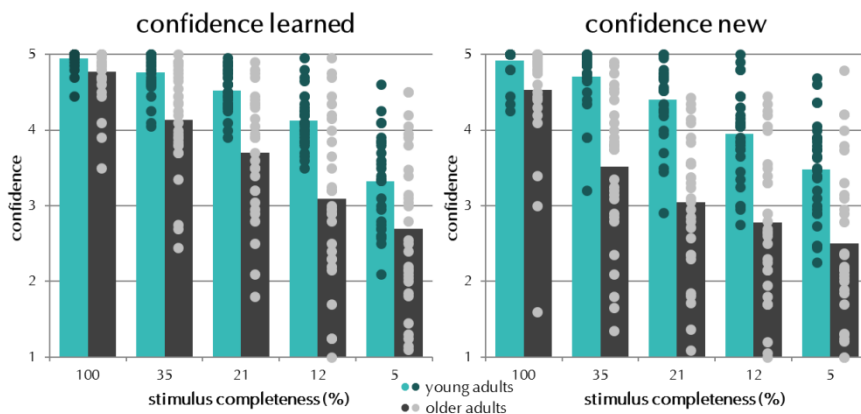
Reaction times followed the profile of performance values as assessed by a three-way mixed ANOVA (age  $\times$  stimulus completeness  $\times$  stimulus type; see Figure 6). Older adults were generally slower than young adults (main effect of age:  $F_{(1,58)} = 25.333$ ,  $p < 0.001$ ), both groups became slower with decreasing stimulus completeness (main effect of stimulus completeness:  $F_{(4,232)} = 37.01$ ,  $p < 0.001$ ), and older adults slowed down more with decreasing information (age  $\times$  stimulus completeness:  $F_{(4,232)} = 4.748$ ,  $p = 0.001$ ). Overall, there was no difference between learned and new stimuli (main effect of stimulus type:  $F_{(1,58)} = 0.008$ ,  $p = 0.93$ ), but interestingly younger adults were faster for new stimuli while older adults were faster for learned stimuli (age  $\times$  stimulus type:  $F_{(1,58)} = 11.867$ ,  $p = 0.001$ ; age  $\times$  stimulus type  $\times$  stimulus completeness:  $F_{(4,232)} = 3.123$ ,  $p = 0.016$ ).



**Figure 6. Ch. 2 - reaction times.** Values for both age groups, separately for learned and new stimuli for the 5 different levels of stimulus completeness (mean).

### 2.3.4. CONFIDENCE RATINGS

The confidence ratings further support the present findings (see Figure 7). A three-way mixed ANOVA (age × stimulus completeness × stimulus type) revealed that older participants were generally less confident than young participants (main effect of age:  $F_{(1,58)} = 33.499$ ,  $p < 0.001$ ). All participants were more confident in their responses when less of the image was masked (main effect of stimulus completeness:  $F_{(4,232)} = 205.047$ ,  $p < 0.001$ ), and were also more confident when responding to learned, relative to new, items (main effect of stimulus type:  $F_{(1,58)} = 26.931$ ,  $p < 0.001$ ). Mirroring task performance, relative to young adults, older participants were less confident in their responses (interaction of age × stimulus completeness:  $F_{(4,232)} = 11.887$ ,  $p < 0.001$ ). Significant interactions (age × stimulus type:  $F_{(1,58)} = 16.733$ ,  $p < 0.001$ ; age × stimulus type × stimulus completeness:  $F_{(4,232)} = 3.58$ ,  $p = 0.007$ ) indicate that the age groups' confidence was differentially affected depending on whether the stimuli were learned or new. As can be seen in Figure 7, older adults were less confident in the identification of new stimuli, whereas young adults were equally confident for learned and new stimuli. This further supports the finding that older adults' performance is adversely affected by unknown stimuli as compared to learned stimuli.



**Figure 7. Ch. 2 - confidence ratings.** Scores for both age groups, separately for learned and new stimuli for the 5 different levels of stimulus completeness (mean). Ratings ranged from 1 ('not at all confident') to 5 ('very confident').

## 2.4. DISCUSSION

We used a novel recognition memory paradigm to assess pattern completion and the impact of cognitive aging on this process. In contrast to previous studies investigating episodic memory processing we have shifted the focus from pattern separation to pattern completion. In the experiment, participants were asked to identify complete or partially masked stimuli, half of which they had learned previously. For both age groups, recognition accuracy was reduced with decreasing stimulus completeness. This effect, however, was more pronounced in older adults, suggesting that pattern completion may be adversely affected by aging. Older adults also showed a response bias toward familiar stimuli, as evidenced by the profile of errors for new items (i.e., a tendency to incorrectly select a familiar item as a response). This behaviour may be the result of an underlying pattern completion bias as suggested by theoretical models of aging (Wilson et al. 2006).

The current paradigm is based on the original computational definition of pattern completion as a recollective process that restores a complete memory trace from partial or degraded input (Marr 1971). Hunsaker and Kesner (2013 p. 40) have suggested that presenting subsets of an original cue could engage pattern completion more independently than degraded versions of it. We have incorporated this idea by manipulating stimulus completeness, in contrast to previous studies that have used altered versions of familiar stimuli (for review, see Yassa and Stark 2011; Stokes et al. 2015). Assuming that our manipulation was successful, the reported recognition memory deficits for previously experienced stimuli in older adults would indicate that pattern completion becomes deficient with age. Paleja and Spaniol (2013) reported similar findings in a task requiring participants to relocate a familiar target in a virtual environment. Older adults performed significantly worse than young adults when fewer cues were available at test, compared to study. This difference, however, was apparent only when no extra-maze cues were available and it remains open whether this could be explained by higher exploration rates of the young participants in comparison to older adults.

In a more general view of age-related memory changes, it is often reported that recognition memory is not as impaired as for example free recall is (Danckert and Craik 2013; Luo and Craik 2008). In contrast, our data demonstrate that recognition memory is impaired in the older age-group in a paradigm like ours and gets significantly more impaired when the retrieval stimuli become less complete. These results can provide a link by illustrating age-related impairments on the spectrum from environmentally aided (recognition) to more self-initiated (recall) (Craik 1983; Luo and Craik 2008). Our finding that older adults were already impaired in the complete conditions (100%) may be explained by the fact that, unlike in many recognition paradigms in the literature, correct responses here demanded exact identification/naming of the stimulus as opposed to old/new/similar or remember/know judgments. We therefore suggest that previous assumptions about recognition memory in aging would need to be revisited in scenarios where the stimuli prompting retrieval are not exactly the same as during learning, or where tasks demand precise identification.

The second major finding of our study – the positive response bias in older adults – shows that this group tended to increasingly choose familiar responses even though they were presented with new stimuli. Thus, new partial information seemed to trigger the recognition of learned items although it was not part of the original cues. We interpret this as a bias toward pattern completion. Our findings would imply that with age even though the process of pattern completion seems to be deficient during actual memory retrieval, it is increasingly initiated despite the new (partial) information not corresponding to a stored memory.

Given its extensive excitatory recurrent connections, region CA3 within the hippocampus has been identified as a likely candidate to execute the auto-associative processing essential for pattern completion (for review, see Hunsaker and Kesner 2013). This theory has very recently received direct empirical evidence from rodent data (Neunuebel and Knierim 2014) showing that CA3 effectively performs pattern

completion on the sensory input it receives from entorhinal cortex and the dentate gyrus. It should be noted that CA3 is not exclusively involved in pattern completion, but has also been found to contribute to other processing like rapid encoding, short-term memory or recall (for review, see Rolls and Kesner 2006). Importantly, aging appears to selectively affect parts of the hippocampal circuit (Smith et al. 2000), because the perforant path degenerates - hence sensory input to CA3 is diminished - while CA3's auto-associative network remains relatively intact. As a result, even when less information is fed forward to CA3, that is, only a subset of neurons is activated, neighbouring CA3 cells can still be co-activated and memory traces be restored. Additionally, there is evidence from the rodent literature showing hyperactivity of CA3 cells in older rats (Wilson et al. 2005), and an age-related decline in cholinergic modulation which reduces inhibition in CA3 (Hasselmo et al. 1995). Theoretical models of hippocampal function therefore predict that older adults should show a bias toward pattern completion (Wilson et al. 2006). Simultaneously, less sensory input to the dentate gyrus is suggested to result in the encoding of less distinct (or pattern separated) memory traces. The conjunction of those degenerative aspects could lead to an increased tendency to reactivate stored memory traces rather than encode new input. Our findings further support this notion, since older adults demonstrated a bias toward familiar stimuli, whereas young adults did not.

As mentioned earlier, the focus of many human studies has been directed towards pattern separation. Yassa and colleagues (2011b) have reported an age-related shift from pattern separation to completion, which correlated with the integrity of the perforant path and the dentate gyrus/CA3 complex. In that study, participants saw pictures of objects with gradually decreasing mnemonic similarity among old and new items, and young adults readily exhibited separation-like BOLD responses for all the similar items while older adults only shifted from completion-like to separation-like activity with a bigger stimulus change. These findings were interpreted as a reduction in pattern separation processes in older adults, based also on behavioural discrimination deficits (i.e. incorrectly identifying a similar item as old). Their use of very similar stimuli might prompt this interpretation; however, we would like to point out that the reported shift might not necessarily stem from impaired pattern separation only. Our results could provide an additional explanation, highlighting the importance of pattern completion. Our stimulus selection did not require strong pattern separation as discussed earlier, because the images here were very different. Behavioural discrimination deficits could therefore also result from a bigger impact of pattern completion processes as compared to pattern separation. However, we cannot rule out the contribution of other processes in our task. Especially during the identification of new stimuli, pattern separation is potentially involved to compare the new item to the retrieved one. The observed response bias in older adults may therefore partly be explained by a failure to separately encode new stimuli, i.e. orthogonalize the new item to the retrieved one.

Our experiment was designed to specifically tap into retrieval processes by introducing a learning criterion during the study phase to ensure equal encoding, yet, it is possible

that young and older adults exhibited differential learning. Given that there is no bias for complete stimuli, however, the observed group differences are likely to result from impaired retrieval processes rather than differential encoding. Along the same lines, we believe one can rule out a perceptual deficit of older adults as a primary cause of the reported bias. Any perceptual influence should follow the linear profile of the completeness manipulation, independent of learned or new stimuli, that is, less visibility should lead to weaker perception. This assumption cannot explain why older adults made more errors identifying new stimuli in the middle completeness levels as compared to the lowest level. Nevertheless, in future studies, it may be worth employing a methodology such as eye-tracking to control for potential encoding or perceptual differences similarly to Molitor and colleagues (2014).

An alternative explanation for a pattern completion bias in older adults is an inability to detect novelty. It has been suggested that the hippocampus acts as a match-mismatch detector that evaluates current sensory input in relation to stored representations to identify novelty (Kumaran and Maguire 2007, 2009). In light of this mechanism, the reported pattern completion bias of older adults could be interpreted as a failure or impairment of novelty detection. Indeed, our results show that older adults are worse at identifying something new. However, Kumaran and Maguire (2007) have argued that hippocampal mismatch signals occur only when new input is very similar to stored memories, and interferes with predictions derived from previous experience. The observed linear performance decline with decreasing stimulus completeness suggests a mechanism signalling the degree of familiarity rather than a pure match-mismatch model as reasoned by Kumaran and Maguire (2009).

Finally, there is a wide literature documenting increased false alarm rates in older adults (Schacter et al. 1997), that is, an age-related increase in judging new items as old. Several reasons have been described to account for this change, including more liberal response criteria, decreasing overall sensitivity, increasing reliance on gist memory, or decreasing item-specific memory. Our data are not consistent with global shifts in sensitivity or response criteria, due to the reported bias curve; i.e. if there was a global criterion shift, false alarm rates should be uniformly distributed across all levels of stimulus completeness, but instead their probability varied. We want to point out that our data are different to some extent, because we did not find an increase in false alarms for complete stimuli, which are the standard material in most previous studies. Related to this, an fMRI study has suggested frontal regions as the origin of higher false alarm rates in older adults as opposed to medial temporal regions, pointing to impaired monitoring (Duarte et al. 2010). However, the univariate analyses employed by the authors might not be sensitive enough to detect activity differences in medial temporal regions, because the average signal intensity in a voxel does not inform about subtle signal variations across voxels and conditions likely to occur during processes like pattern completion.

One major limitation of our study is that we cannot unravel the underlying neural processes and identify the involvement of different hippocampal subfields with a behavioural experiment like this. It may well be that pattern separation in the dentate

gyrus is also impaired during encoding or even retrieval, while there are pattern completion deficits in CA3, and that the conjunction of these alleged processes produced the observed behaviour. It remains a challenge for future neuroimaging research to disentangle these different processes, the involved brain areas and the resulting behavioural responses.

To summarize, we have used a novel recognition memory paradigm designed to target pattern completion processes by manipulating stimulus completeness. On the one hand, we demonstrated age-related recognition memory deficits for learned items, suggesting an underlying deficit in pattern completion. On the other hand, we showed a bias in older adults toward familiar responses during the identification of new items, strongly suggesting increased initiation of pattern completion processes even though the trigger stimuli should not have a corresponding memory trace. These findings are in line with predictions derived from theoretical models based on the literature about hippocampal circuitry. Our results provide more detailed insights into recognition memory deficits reported in older adults, and they serve as a starting point for further investigations to shed light on the underlying neural processes.

## 2.5. CONTRIBUTIONS

Matthias Stangl, Lorelei Howard and Thomas Wolbers helped conceptualize the study. Christin Ruß and Franziska Schulze recruited and tested most of the participants. Jonathan Shine made valuable comments on the manuscript before publication.

## 3. EYE-MOVEMENTS IN THE MIC

### How do eye-movements contribute to age-related recognition memory differences assessed by the MIC?

#### 3.1. INTRODUCTION

The paradigm presented in chapter 2.2 (now called Memory Image Completion) was carefully designed to more directly tap into pattern completion processes rather than being tested as a by-product of a pattern separation task. However, contributions of other sources could not be ruled out by virtue of a behavioural task only. Therefore, here, I used eye-tracking to eliminate further confounds and identify general eye movements associated with this type of task and their relation to cognitive aging.

Viewing behaviour in the context of memory research can be used to indicate memory-related processing. For example, higher fixation numbers during encoding have been suggested to reflect increased accumulation of information leading to better memory representations (Pertzov et al. 2009) and in turn resulting in increased memory performance (Loftus 1972). Similarly, during retrieval longer viewing durations have been associated with previously studied faces as opposed to very similar new faces also independent of explicit recognition (Hannula et al. 2012), and this relationship coincided with hippocampal activity (Hannula and Ranganath 2009). Furthermore, elevated fixation rates during retrieval corresponded to recollection rather than familiarity (Kafkas and Montaldi 2012). To my knowledge, there are no memory-related eye-tracking studies investigating the effects of aging. However, more general visual changes that are tied to aging include slower saccade velocity and saccadic reaction times (Moschner and Baloh 1994), as well as lower saccade frequencies and amplitude (Dowiasch et al. 2015).

Next to that, eye-tracking has also been used to question the process-purity of pattern separation and completion tasks (Molitor et al. 2014). Specifically, age-related reduced correct rejections and increased false alarms to lures had previously been attributed to a shift from pattern separation to pattern completion as a retrieval based mechanism (Yassa et al. 2011b), i.e. very similar stimuli were more likely to be identified as old (pattern completion bias) than as similar (pattern separation failure). However, the authors of the eye-tracking study (Molitor et al. 2014) found fewer fixations on these items during encoding, suggesting that a purely retrieval-based process is unlikely or at least insufficient to explain the observed behavioural bias.

Even though my paradigm does not use incidental encoding and already tries to accommodate potential encoding differences through a learning criterion, I wanted to account for possible perceptual encoding differences and ascertain if viewing patterns

had any explanatory power over age-related differential retrieval. Thus, I tested the recognition memory paradigm described in 2.2 again on young and older adults with simultaneous eye-movement recordings.

## 3.2. METHODS

### 3.2.1. SUBJECTS

Twenty-six young (21-35 years old; 13 females; 6 wore glasses, 7 wore contact lenses) and 24 older adults (63-77 years old; 12 females, 22 wore glasses) were included in the study. They were recruited from existing databases at the DZNE in Magdeburg, and underwent several neuropsychological tests and health assessments prior to the experiment: health questionnaire (Diersch 2013), MoCA (Nasreddine et al. 2005), multiple choice word test (MWT-B; Lehrl et al. 1995), digit symbol substitution test (DSST; Wechsler 2008), Rey-Osterrieth Complex Figure (ROCF) Test as copy and 30 minutes delayed recall (DR) test (Rey 1941; Corwin and Bylsma 1993); their visual dominance was assessed, and subjective (using German grading system 1 (very good) – 6 (insufficient)) and objective eyesight (visual acuity determined on a pocketcard test). Participants who scored less than 23 on the MoCA (N = 2; Luis et al. 2009), who had chronic psychological or neurological problems (N = 1), who had insufficient eyesight (N = 3), or whose eyes could not be tracked due to their glasses (N = 4; e.g. reflection, small frame, dark frame), had been excluded from the study. For the remaining participants, see Table 2 for descriptive neuropsychological data and answers to the health questionnaire, included to better describe the tested population. Informed written consent was obtained before the experiment, and the study received approval from the Ethics Committee of the University of Magdeburg. All participants received monetary compensation of 6.50€/h.

**Table 2. Ch. 3 - health questionnaire and neuropsychological data.**

age group	school (years)	higher education (years)	MoCA score	MWT-B		digit symbol substitution		ROCF test		vision	
				raw	%	raw	score	copy	30 min DR	sub-jective	ob-jective
young	12.83 (0.89)	4.42 (2.81)	28.36 (0.93)	30.72 (2.66)	75.98 (16.47)	88.58 (17.51)	12.0 (3.16)	35.16 (1.44)	25.06 (5.82)	1.71 (0.75)	0.71 (0.1)
old	10.63 (1.34)	5.09 (3.22)	26.96 (1.99)	32.96 (1.51)	91.84 (8.45)	56.88 (10.84)	11.13 (1.83)	34.35 (2.37)	18.02 (4.3)	2.04 (0.51)	0.63 (0.12)

Mean values (standard deviation). MoCA – Montreal Cognitive Assessment; MWT-B – multiple choice word test; ROCF – Rey-Osterrieth complex figure; DR – delayed recall

### 3.2.2. MATERIALS & PROCEDURE

This experiment was an extension of the task presented above (2.2) with simultaneous eye-tracking. That is, the behavioural paradigm and procedure were exactly the same as in PatComp and are described above. Additionally, eye movements were recorded during image presentation in all parts of the experiment. Participants were seated 50 centimetres in front of the display monitor. At the beginning of each phase of the experiment, eye position was calibrated with a 9-fold grid of fixation points. Participants were asked to refrain from looking down at the keyboard when responding,



because head movements could cause slight shifts of the eye-tracker. However, as participants had to perform several button presses with at least 5 response options per trial, this could not be prevented entirely. Therefore, every trial started with a drift correction prior to image presentation to match calibration, in which participants had to fixate a small white circle in the middle of the screen. If drift correction failed (drift > 5° visual angle), tracking of eye position was adjusted by recalibration. During the test phase, eye position was recalibrated every 70 trials by default. The participants' dominant eye was always recorded (32 right, 18 left), except for two subjects, for whom eyes were switched after half of the experiment due to a recalibration failure.

---

### 3.2.3. EYE-TRACKING ACQUISITION AND ANALYSIS

Eye movements were recorded with a head-mounted EyeLink II tracker (SR Research, Ontario, Canada) at a sampling rate of 500 Hertz. Stimuli were displayed on a 15 inch computer screen (1024 × 768 pixel resolution, 60 Hertz refresh rate). The experiment was programmed in Matlab 2013a with the Psychtoolbox Add-on to integrate the eye-tracker. Eye movements were recorded for each participant, and saccade, blink, and fixation data were calculated as follows. Blinks were defined as missing pupil data over three consecutive samples. Saccades were eye movements greater than 0.1° visual angle, and faster than 30°/s velocity and 8000°/s<sup>2</sup> acceleration. Fixations consisted of all other recordings. Eye-tracking analyses focused on number and duration of fixations during the 2 second image presentation at study and test, and their proportion in pre-defined regions of interest (ROIs). ROIs were equivalent to the inverse masking grid used to manipulate stimulus completeness, i.e. they consisted of all the areas in the image that were still visible through the mask. Eye-movement data of two young participants were excluded for the first half of the study phase, because of a calibration error.

---

## 3.3. RESULTS

---

### 3.3.1. BEHAVIOURAL RESULTS

All analysis steps were identical to the previous chapter (2.3), and replicated the main results.

---

#### 3.3.1.1. ACCURACY

The following results were obtained from a three-way mixed ANOVA with factors age, stimulus completeness, and stimulus type. Young participants performed better than older participants (main effect of age:  $F_{(1,48)} = 142.086$ ,  $p < 0.001$ ). Reduced stimulus completeness resulted in less accurate performance (main effect of stimulus completeness:  $F_{(4,192)} = 153.102$ ,  $p < 0.001$ ), and this decrease was more pronounced in older adults (age × stimulus completeness:  $F_{(4,192)} = 33.582$ ,  $p < 0.001$ ). Performance per stimulus type did not differ overall (main effect of stimulus type:  $F_{(1,48)} = 0.568$ ,  $p = 0.455$ ), but interactions revealed that older participants in fact were less accurate for new stimuli as compared to learned stimuli (age × stimulus type:  $F_{(1,48)} = 5.408$ ,  $p = 0.024$ ; age × stimulus type × stimulus completeness:  $F_{(4,192)} = 14.582$ ,  $p < 0.001$ ). Post-

hoc independent t-tests showed age group differences in performance across all levels of stimulus completeness for both learned and new stimuli (after Holm-Bonferroni multiple comparisons correction; all  $p < 0.01$ ; level 100%:  $t_{learned(48)} = 2.887$ ,  $t_{new(48)} = 2.936$ , level 35%:  $t_{learned(48)} = 4.667$ ,  $t_{new(48)} = 11.045$ , level 21%:  $t_{learned(48)} = 5.676$ ,  $t_{new(48)} = 12.418$ , level 12%:  $t_{learned(48)} = 6.508$ ,  $t_{new(48)} = 9.196$ , level 5%:  $t_{learned(45.782)} = 4.747$ ,  $t_{new(48)} = 4.024$ ). In summary, both age groups' recognition ability decreased with reduced stimulus completeness; older adults' recognition ability was impaired in comparison to young adults', and especially so for new stimuli.

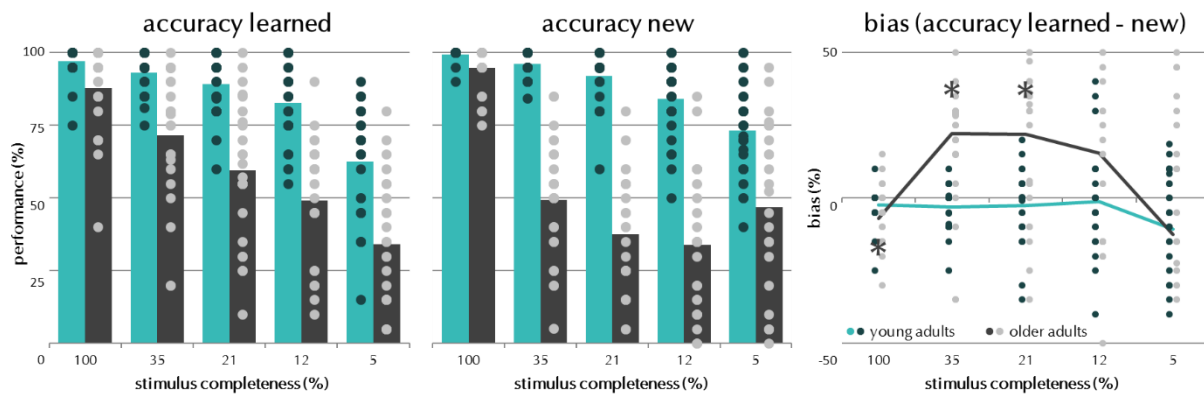
---

### 3.3.1.2. RESPONSE BIAS

Again, I specifically looked at potential response biases, by subtracting individual accuracy scores of new stimuli from those of learned stimuli (see chapter 2.2.4 for details). Positive scores are indicative of a bias towards pattern completion, i.e. better performance for to-be-retrieved (learned) stimuli; and respectively, negative scores indicate a bias towards pattern separation, i.e. better performance for to-be-encoded (new) stimuli.

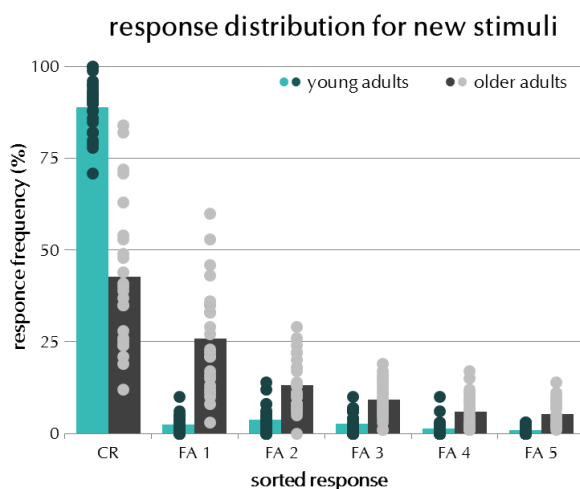
A mixed ANOVA (age  $\times$  stimulus completeness) revealed that stimulus completeness influenced the bias scores (main effect of stimulus completeness:  $F_{(4,192)} = 14.582$ ,  $p < 0.001$ ), and older adults showed a more positive bias than young adults (see Figure 8, right; main effect of age:  $F_{(1,48)} = 5.408$ ,  $p = 0.024$ ). A two-way interaction indicated that stimulus completeness differently affected the bias scores dependent on the age group (age  $\times$  stimulus completeness:  $F_{(4,192)} = 8.053$ ,  $p < 0.001$ ). Follow-up independent t-tests demonstrated significant between-group differences for the middle three completeness levels, but after Holm-Bonferroni multiple comparisons correction the 12% level did not reach the necessary significance level (level 35%:  $t_{(48)} = -4.68$ ,  $p < 0.001$ , level 21%:  $t_{(48)} = -4.042$ ,  $p < 0.001$ ; level 12%:  $t_{(48)} = -2.107$ ,  $p = 0.04$ ).

I also tested at which levels of stimulus completeness the scores established a bias. Therefore, five one-sample t-tests of the group average bias scores against 0 revealed that older adults showed a positive bias for completeness levels 35% and 21%, indicative of a pattern completion bias, and a negative bias for level 100% only just crossing the threshold after Holm-Bonferroni multiple comparisons corrections (level 100%:  $t_{(23)} = -2.416$ ,  $p = 0.024$ ; level 35%:  $t_{(23)} = 4.149$ ,  $p < 0.001$ , level 21%:  $t_{(23)} = 3.819$ ,  $p = 0.001$ ; level 12%:  $t_{(23)} = 2.081$ ,  $p = 0.049$ ; level 5%:  $t_{(23)} = -1.392$ ,  $p = 0.177$ ). Young participants showed no bias at all.



**Figure 8. Ch. 3 - performance and bias measures.** Left, performance for both age groups, separately for learned and new stimuli for the 5 different levels of stimulus completeness (mean); right, bias measure (see 2.2.4 for a detailed explanation) - difference in accuracy scores for learned minus new stimuli calculated separately for each participant (mean); positive values indicate a bias toward pattern completion, significant differences from 0 are indicated with \* separately for each age group as indicated by colour.

Again, I looked at the distribution of false alarm options for new stimuli. If participants completed towards the stimulus perceived as most similar, rather than simply guessed more, one false response option should have been chosen over all other false responses. Therefore, I calculated the group-average frequencies of all false choice options and sorted them from most (FA 1) to least (FA 5) chosen option for each new stimulus, and averaged across all new stimuli afterwards. A  $\chi^2$ -test of goodness-of-fit on the 5 false alarm options ( $\chi^2 = 429.009$ ,  $df = 4$ ,  $p < 0.001$ ) revealed that participants indeed chose one particular false response option most often (FA 1) and did not simply guess more. I also checked whether there was one dominant option for each stimulus, by contrasting the most frequent false alarm against the average of the other false alarms per stimulus (stimulus 'office':  $\chi^2 = 108.938$ ,  $df = 1$ ,  $p < 0.001$ ; stimulus 'class room':  $\chi^2 = 24.0$ ,  $df = 1$ ,  $p < 0.001$ ; stimulus 'restaurant':  $\chi^2 = 23.343$ ,  $df = 1$ ,  $p < 0.001$ ; stimulus 'locker room':  $\chi^2 = 11.172$ ,  $df = 1$ ,  $p = 0.001$ ; stimulus 'living room':  $\chi^2 = 5.851$ ,  $df = 1$ ,  $p = 0.016$ ). This indicates that especially older participants completed towards the stimulus perceived as most similar.



**Figure 9. Ch. 3 - response distribution for new stimuli.** Left, responses are depicted over the 6 possible choice options (i.e., 'none of these' as correct rejections - CR, and the false alarms sorted according to frequency - FA 1-5; mean) showing that older adults chose one particular false response option most often (FA 1) rather than guess more overall, which would lead to similar frequencies for all 5 response options.

For the sake of completeness, I present false alarm rates for learned stimuli, but there are not sufficient data for an adequate statistical analysis. Descriptively, older adults

have higher false alarm rates than young adults, though here the false alarms do not surpass misses until the least complete versions of the stimuli.

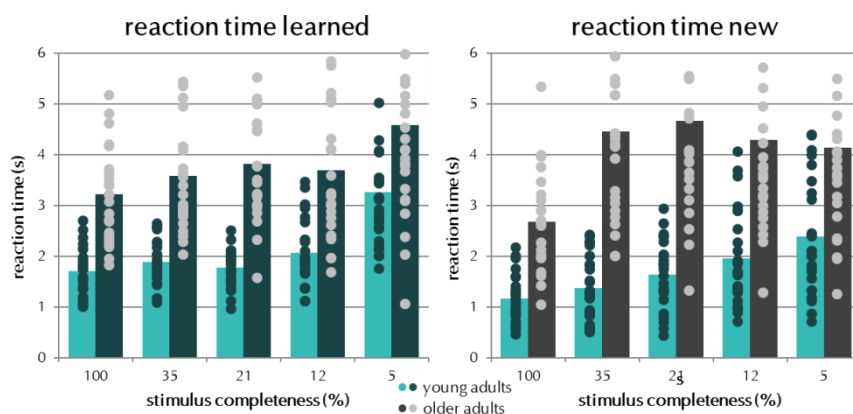
**Table 3. Ch. 3 - false alarm rates for learned stimuli.**

stimulus completeness	false alarms - mean (SE)	
	young adults	older adults
100%	0.08 (0.01)	0.18 (0.03)
35%	0.11 (0.01)	0.30 (0.04)
21%	0.14 (0.02)	0.42 (0.05)
12%	0.21 (0.02)	0.51 (0.04)
5%	0.37 (0.04)	0.66 (0.04)

False alarms and misses add up to 1, so that values higher than 0.5 indicate more false alarms, and values lower than 0.5 indicate more misses; values do not comprise the data of all participants since not all of them made errors for each completeness level.

### 3.3.1.3. REACTION TIMES

Reaction times were also analysed by a three-way mixed ANOVA (age × stimulus completeness × stimulus type), and mirrored task performance (see Figure 10). Older adults were generally slower than young adults (main effect of age:  $F_{(1,48)} = 45.044, p < 0.001$ ), both groups became slower with decreasing stimulus completeness (main effect of stimulus completeness:  $F_{(4,192)} = 35.993, p < 0.001$ ), but this trend was more consistent in younger adults (age × stimulus completeness:  $F_{(4,192)} = 7.899, p = 0.001$ ). Overall, there was no difference between learned and new stimuli (main effect of stimulus type:  $F_{(1,48)} = 1.154, p = 0.288$ ), but younger adults were faster for new stimuli while older adults were faster for learned stimuli, respectively compared to the other stimulus type (age × stimulus type:  $F_{(1,48)} = 19.992, p < 0.001$ ; age × stimulus type × stimulus completeness:  $F_{(4,192)} = 14.215, p < 0.001$ ).

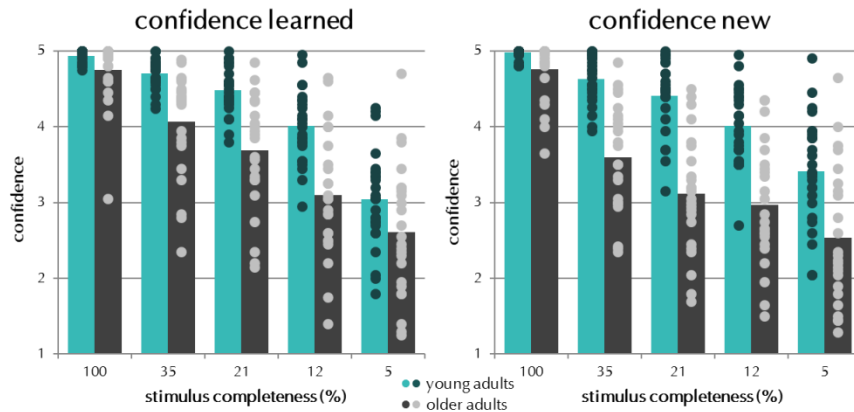


**Figure 10. Ch. 3 - reaction times.** Values for both age groups, separately for learned and new stimuli for the 5 different levels of stimulus completeness (mean).

### 3.3.1.4. CONFIDENCE RATINGS

The confidence ratings further support the present findings (see Figure 11). A three-way mixed ANOVA (age × stimulus completeness × stimulus type) revealed that older participants were generally less confident than young participants (main effect of age:  $F_{(1,48)} = 16.76, p < 0.001$ ). All participants were more confident in their responses when less of the image was masked (main effect of stimulus completeness:  $F_{(4,192)} = 162.451, p < 0.001$ ), and were also more confident when responding to learned, relative to new,

items (main effect of stimulus type:  $F_{(1,48)} = 5.777, p = 0.02$ ). Following the profile of performance values, relative to young adults, older participants were less confident in their responses (interaction of age  $\times$  stimulus completeness:  $F_{(4,192)} = 10.351, p < 0.001$ ). Significant interactions (age  $\times$  stimulus type:  $F_{(1,48)} = 13.254, p = 0.001$ ; age  $\times$  stimulus type  $\times$  stimulus completeness:  $F_{(4,192)} = 16.676, p < 0.001$ ) indicate that older adults were less confident in the identification of new stimuli, whereas young adults were equally confident for learned and new stimuli. This adds to the evidence that older adults have a specific impairment when presented with unknown stimuli as compared to learned stimuli.



**Figure 11. Ch. 3 - confidence ratings.** Scores for both age groups, separately for learned and new stimuli for the 5 different levels of stimulus completeness (mean). Ratings ranged from 1 ('not at all confident') to 5 ('very confident').

### 3.3.2. EYETRACKING RESULTS

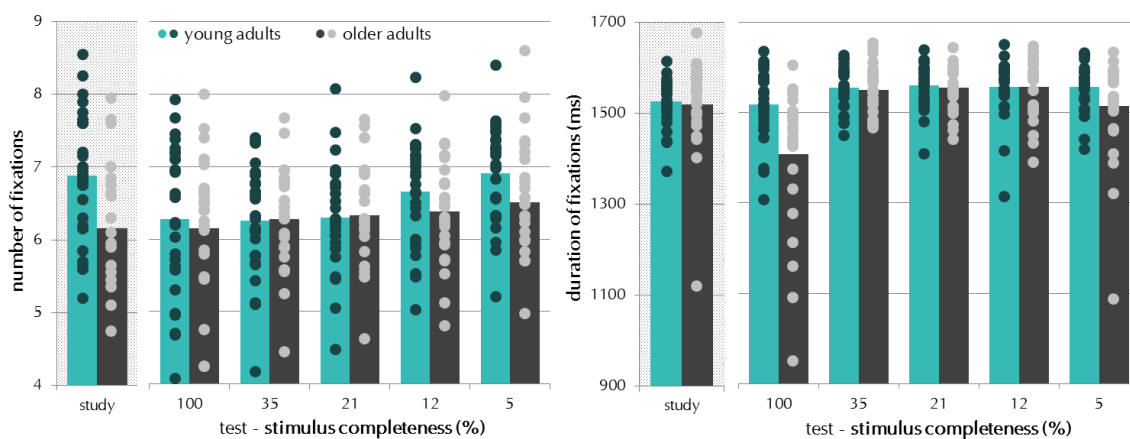
#### 3.3.2.1. TOTAL FIXATIONS

First, I looked at the overall number of fixations in a similar way as for behavioural accuracy. I subjected the fixation numbers to a three-way mixed ANOVA with factors age (young, old), stimulus completeness (100%, 35%, 21%, 12%, 5%) and stimulus type (learned, new). Interestingly, young and older adults did not differ in how much they fixated on the images (main effect of age:  $F_{(1,48)} = 0.514, p = 0.477$ ), nor was there a difference between learned and new stimuli (main effect of stimulus type:  $F_{(1,48)} = 0.238, p = 0.628$ ). However, the participants fixated slightly more when less of the image was visible (main effect of stimulus completeness:  $F_{(4,192)} = 11.128, p < 0.001$ ).

Next, the sum of durations of the fixations was analysed analogously with a three-way mixed ANOVA revealing the same types of effects. There is an inverse relationship between the number and a single duration of fixations, i.e. the longer participants spent fixating on one location, the lower the number of fixations was and vice versa. This is somewhat logical because participants only had 2 seconds to look at the image resulting in a trade-off between how many locations to look at and how closely to look at each of them. Therefore, I used the sum of durations for all fixations rather than mean duration, as they present a more meaningful measure. While there were no main effects for age or stimulus type (main effect of age:  $F_{(1,48)} = 3.189, p = 0.08$ ; main effect of stimulus type:  $F_{(1,48)} = 0.714, p = 0.402$ ), the masking levels differently affected fixation durations (main effect of stimulus completeness:  $F_{(4,192)} = 24.231, p < 0.001$ ), and here, also interacted with age (age  $\times$  stimulus completeness:  $F_{(4,192)} = 8.292, p < 0.001$ ). More

specifically, older adults fixated only shortly on the full stimuli but spend more time fixating the masked stimuli with durations similar to that of young adults whose fixation duration did not differ across masking levels (see right panel in Figure 12). Given the appearance of a masked stimulus, one may argue that the observed viewing pattern is a useful strategy for this task. That is, with decreased stimulus completeness there is less to see at each fixation rendering it necessary to shift fixations more often to obtain a similar amount of information as for the unmasked stimuli.

To investigate encoding effects, I also looked at fixations during learning. There were no differences in neither fixation numbers nor durations between groups (fixation number:  $t_{(34,014)} = 1.869$ ,  $p = 0.07$ ; fixation duration:  $t_{(33,714)} = 0.814$ ,  $p = 0.762$ ; see shaded plots in Figure 12). This is taken to suggest no substantial age-related differences in encoding.



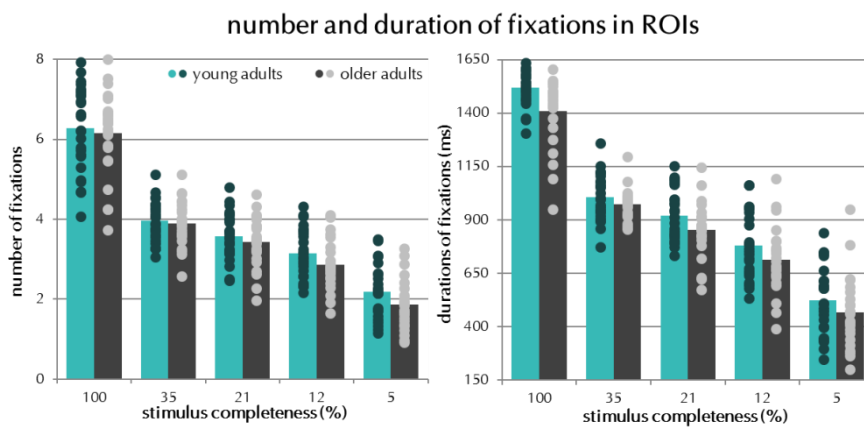
**Figure 12. Ch. 3 - eye-tracking data for the whole images at study and test.** Viewing behaviour during learning (study) is shaded. Left, number of fixations for both age groups, combined over learned and new stimuli for the 5 different levels of stimulus completeness (mean  $\pm$  SE); right, mean duration of fixations in ms for both age groups and separately for the 5 stimulus completeness levels (mean  $\pm$  SE).

### 3.3.2.2. FIXATIONS IN ROIS

Next, I inspected target regions to identify where participants looked. Therefore, I examined fixations in predefined ROIs. As explained in the methods (3.2.2), the stimuli were masked in different levels, so that only parts of the image were visible. I wanted to investigate whether participants would look at the relevant positions, i.e. the parts of the image that still carried information. Thus, ROIs constituted all the unmasked portions of the image.

From there, fixations within the ROIs were analysed with a three-way mixed ANOVA. Again, young and older adults did not differ overall in their number of fixations on ROIs (main effect of age:  $F_{(1,48)} = 1.472$ ,  $p = 0.231$ ), nor were there differences between learned and new stimuli (main effect of stimulus type:  $F_{(1,48)} = 0.769$ ,  $p = 0.385$ ). Not surprisingly, both groups had less fixations on less complete images (main effect of stimulus completeness:  $F_{(4,192)} = 458.846$ ,  $p < 0.001$ ). That was to be expected, because the ROIs got smaller with increasing masking levels. The same was true for fixation durations, i.e. all participants fixated shorter on smaller ROIs and spend more time fixating on the more complete stimuli (see middle panel in Figure 13; main effect of stimulus completeness:  $F_{(4,192)} = 729.865$ ,  $p < 0.001$ ), and there were no differences

between learned and new stimuli (main effect of stimulus type:  $F_{(1,48)} = 3.749$ ,  $p = 0.059$ ). However, here younger adults had slightly longer fixation durations than older adults (main effect of age:  $F_{(1,48)} = 4.896$ ,  $p = 0.032$ ).



**Figure 13. Ch. 3 - eye-tracking data for the regions-of-interest (ROIs).** ROIs were the parts of the image still visible through the masks. Data are collapsed over the stimulus types and separate for the 5 different levels of stimulus completeness for both age groups (mean  $\pm$  SE). Left, number of fixations; right, sum of durations of fixations in ms.

### 3.3.3. SUMMARY OF THE RESULTS

The findings suggest that viewing behaviour does not drive age-related performance differences on the task. Specifically, young adults were similarly accurate for both types of stimuli, while older adults performed worse for new stimuli. Additionally, younger adults responded considerably faster to new stimuli than to learned stimuli, whereas it was the other way around for older adults. The only group difference in the eye-tracking data were slightly increased viewing durations of younger adults within ROIs but independent of whether they were learned or new.

## 3.4. DISCUSSION

To summarize, I could replicate the behavioural findings of the first study (2.3), that is, recognition memory declined with reduced stimulus completeness, more so in aging, and older adults were biased towards pattern completion, i.e. they chose familiar responses over new ones. Associated eye-tracking data only showed slightly shorter ROI fixation durations of older adults but with no distinction between the different types of stimuli or any other condition. While this particular difference may contribute to the overall performance reduction of older adults, it cannot account for the more specific impairment in the recognition of new items. Thus, the observed differential recognition memory effects cannot be explained by the corresponding eye-movements, lending support for the pattern completion hypothesis.

Specifically, during study older adults showed similar viewing patterns to that of young adults, indicating that stimuli were physically encoded just as well. The previous literature on memory-related viewing behaviour has linked increased fixation numbers during encoding to better memory retrieval (Loftus 1972). This would suggest that given their impaired retrieval, older adults should show lower fixation numbers at encoding, which was ruled out in this experiment; rather, although there were no differences in viewing patterns during encoding, older adults performed worse during retrieval. Given that the study mentioned above did only report a descriptive relationship between these

measures, I can only assume that differences picked up by eye-tracking may depend more on the specific tasks. Consequently, here, performance differences cannot be explained by differential eye movements during encoding suggesting that another process is involved.

Further on, fixation durations during retrieval were slightly longer in the younger age group potentially accounting for their overall recognition advantage. This idea receives support from the previous findings that longer fixations during retrieval would code for a prior exposure indicating recognition (Hannula et al. 2012; Molitor et al. 2014). However, reports in the literature also show that saccade velocity and reaction times decrease with age (Moschner and Baloh 1994) in line with a general slowing of processing speed (O'Shea et al. 2016). Based on this, shorter fixation durations in older adults could be explained by their increased saccade durations, i.e. eye-movements between fixations. Thus, shorter fixation durations might well account for recognition memory deficits, however, they do not provide a reliable index of the specificity of the underlying process, be it processing speed or memory decline. In further support, the main age-related disparity in this task was observed in the difference in performance for learned and new stimuli which was not present in the eye-tracking data. Hence, while older adults made more errors for new stimuli than for learned ones, their eye-movements were not differentiable between the two stimulus types. Again, the specificity of their errors, i.e. more often picking a specific familiar response instead of the 'none of these' option in identifying new stimuli, and the absence of a corresponding viewing pattern, lends further credibility to the theory that the age differences observed with this task are a result of a pattern completion bias. It is also worth noting that the behavioural findings could almost completely be replicated from their first implementation as reported in chapter 2.3. The only deviation was a non-significant group difference in the bias score for 12% images, but the overall shape of the curve was the same, and so were all other results concerning accuracies, reaction times, confidence or false alarm distributions. This provides compelling evidence that the task in its current form can be reliably applied and the obtained results are not coincidental, thus, it can be used to pick up fine-grained age-related recognition memory differences in the scheme of pattern completion.

### 3.5. CONTRIBUTIONS

Thomas Wolbers helped conceptualize the study and discussed analyses. Henrike Raith recruited most of the participants, ran the neuropsychology tests and acquired behavioural and eye-tracking data as part of a lab rotation. Franziska Schulze and Patrick Hauff acquired some additional eye-tracking data sets, and Franziska Schulze and Anica Luther ran the corresponding neuropsychology tests. Martin Riemer programmed the eye-tracking parameter extraction and I discussed potential analysis steps and interpretations with him.



## 4. DG LESIONS AND THE MIC

### MIC performance assessed in a patient with bilateral DG lesions

The study presented in this chapter has been adjusted to fit the structure and scope of this thesis; the text has been paraphrased from its published version, and the figures were adapted accordingly:

Baker S, Vieweg P, Gao F, Gilboa A, Wolbers T, Black SE, Rosenbaum RS (2016). The Human Dentate Gyrus Plays a Necessary Role in Discriminating New Memories. *Current Biology* 26: 2629–2634.

#### 4.1. INTRODUCTION

So far, I have addressed pattern completion and separation in a merely theoretical and surrogate fashion. That is, I have attributed the computationally defined and neurally attested processes to behaviour, be it performance or viewing patterns. However, none of this is direct evidence that the neural processes in question along with the putative structures are indeed underlying the observed behaviour. Here, I present findings from a collaborative case study carried out in Toronto (Baker et al. 2016) investigating a 54-year-old man, B.L., with selective bilateral DG lesions (Rosenbaum et al. 2014). In the following, I have restructured and paraphrased the published article.

As explained in the introduction (1.3), pattern separation and completion are thought to rely on distinct hippocampal subfields (Treves and Rolls 1994). More specifically, pattern separation has been linked to DG, and pattern completion to CA3 in computational models (O'Reilly and McClelland 1994; Rolls 2016), and rodent studies (Neunuebel and Knierim 2014; Kesner et al. 2016; McHugh et al. 2007). Human fMRI studies have so far pinpointed the segregation at CA1 being responsible for pattern completion and the CA3/DG complex for pattern separation, mainly owed to insufficient resolution (Lacy et al. 2011; Bakker et al. 2008). But there is recent evidence emphasizing a key role of DG only in pattern separation (Berron et al. 2016). Given the limitations of non-invasive studies in humans, studying B.L. presents an exceptional opportunity to disentangle the involvement of DG in specific memory mechanisms.

Consequentially, the most common behavioural task to approximate pattern separation, the MST (Stark et al. 2015), was administered to B.L. Herein, discrimination abilities are tested after incidental learning. To be exact, participants have to classify simple objects as 'old', 'new' or 'similar', with the main focus on similar items. That is, the MST assesses how well participants can identify very similar objects as such rather than classifying them as old. This ability is based on pattern separation, where very similar representations are orthogonalized in DG to allow for differentiation and therefore correct recognition. Given B.L.'s reduced DG volume, we hypothesized that he should

show a deficit in pattern separation, manifesting in a failure to correctly classify very similar items in the MST.

Additionally, the MIC (as described in chapter 2.2) was used to assess pattern completion. The relative intactness of B.L.'s CA3 would suggest that B.L. should not present any deficits in this process. However, the fact that DG projects onto CA3 (Smith et al. 2000; Small 2014) might lead to some sort of impairment. Specifically, it was assumed that B.L. would demonstrate a bias towards pattern completion, as the CA3 auto-associative function may be strengthened similarly as in aging models (Wilson et al. 2006). This would lead to increased retrieval of already stored information to the detriment of encoding new events, manifesting in the discrepancy between learned and new stimuli on the MIC.

## 4.2. METHODS

### 4.2.1. SUBJECTS

This was a case study investigating a 54-year-old man, B.L., with selective bilateral DG lesions, resulting from an electrical injury and cardiac arrest in 1985 (Rosenbaum et al. 2014). Volumetric assessment of B.L.'s hippocampus revealed a 50% smaller DG compared to that of 119 age-matched controls (Mueller and Weiner 2009). It cannot be excluded that a proportion of the lesion also extends into CA3, however visual inspection on high-resolution 3T MRI scans showed that the majority of the volume loss is in fact within DG (Baker et al. 2016). Standard neuropsychological testing revealed that B.L. had mildly impaired anterograde memory and moderately impaired retrograde memory, but was otherwise cognitively unimpaired (Baker et al. 2016).

Additionally, a healthy age-matched control group was recruited from the Baycrest Health Sciences participant pool. Twenty (mean age = 52 years; 10 males) and 19 participants (mean age = 51 years; 13 males) respectively, were tested in two different experiments; 13 of which were common to both experiments. Controls had no history of psychiatric or neurological illness. They received monetary compensation of \$15/h for their participation. Informed consent was obtained in accordance with the Ethics Review Boards at York University and Baycrest, and conforms to the standards of the Canadian Tri-Council Research Ethics guidelines.

### 4.2.2. MATERIALS & PROCEDURE

The experimental protocol included two behavioural tasks: (1) the Memory Image Completion task (MIC; see chapter 2.2) was administered to B.L. and 19 controls; and (2) the Mnemonic Similarity Task (MST; Stark et al. 2013) was administered to B.L. (in two versions) and 20 controls. The MIC followed exactly the same procedures described above (see 2.2). For the MST, participants first saw 128 colour images of everyday objects (e.g. apple, hair brush) and had to judge whether they were indoor or outdoor objects. The stimuli were presented for 2 s each, separated by a 0.5 s inter-stimulus interval (ISI). Afterwards, participants had to perform a surprise recognition memory task on 192 stimuli, indicating whether the stimuli were old, new or similar to the previously seen items. Those stimuli could either be learned items (64 targets),

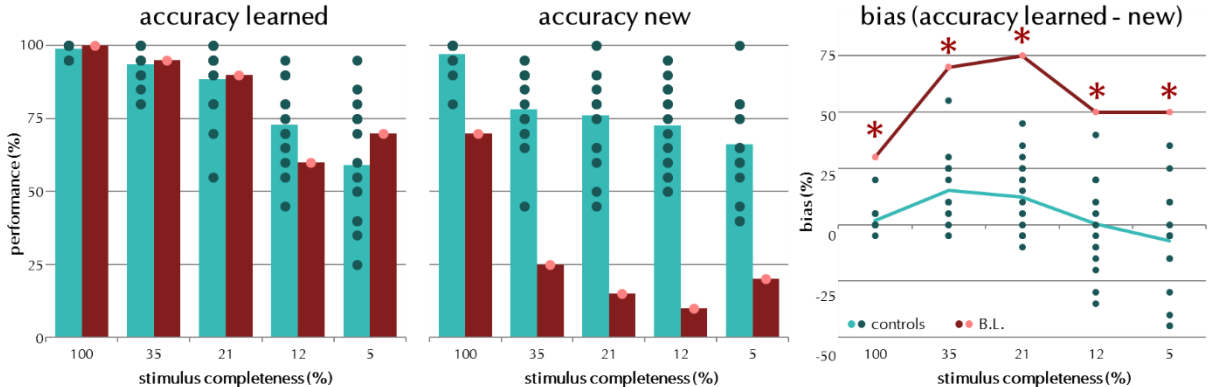
completely new items (64 foils), or similar items (64 lures), and were again presented for 2 s each, separated by 0.5 s ISIs. Performance was assessed through the Lure Discrimination Index (LDI) score (Stark et al. 2013). For each participant, the difference between the rate of similar responses given to lure items minus the rate of similar responses given to foils was calculated, and then averaged across participants to assess the group's pattern separation ability.

### 4.3. RESULTS

Here, single case statistics were employed using Crawford and Howell's modified t-test (Crawford and Garthwaite 2002; Crawford and Howell 1998). Firstly, assessing pattern completion, analyses of the MIC performance revealed that B.L. was as good as controls at identifying learned stimuli, but considerably worse at identifying new stimuli (see left panels of Figure 14). Collapsed over all masking levels, B.L. responded correctly only for 28% of the new stimuli, while controls performance was at 79% ( $t_{(18)} = -2.92, p = 0.01$ ). Again, this difference between learned and new stimuli was further investigated with a bias score by subtracting accuracy values for new stimuli from that of learned stimuli (for details, see 2.2.4). B.L.'s bias scores were higher than that of controls and manifested at all completeness levels ( $p < 0.05$ ; see right panel of Figure 14). Therefore, B.L. seems to have a very strong tendency to pattern complete.

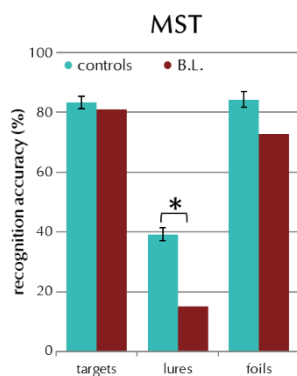
Just as in previous versions of the task, it was checked whether B.L. only guessed more than controls or chose specific wrong responses. When analysing false alarms for new stimuli, B.L. followed the same pattern as older adults in the previous versions of this task (see 2.3.2 and 3.3.1.2), i.e. he chose one specific false alarm option over the others (e.g. he thought the new stimulus 'office' was the learned 'library').

In contrast to controls and all other populations tested previously with this task (see 2.3.1 and 3.3.1.1), B.L. also showed a deficit for complete new stimuli. This points to a failure to dissociate new input from stored representations, which could be interpreted as a failure to pattern separate. However, when contrasting the average bias of all masked stimuli with the bias for full stimuli, this difference was significantly higher in B.L. than in controls ( $p = 0.03$ ), suggesting that B.L.'s bias to pattern complete was stronger than his failure to pattern separate.

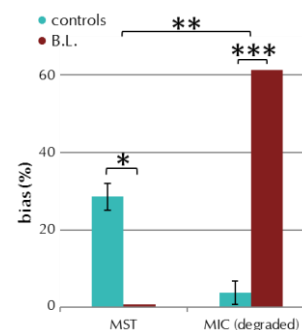


**Figure 14. Ch. 4 - Performance and bias measures.** Left, performance for B.L.(red) and controls (turquoise), separately for learned and new stimuli for the 5 different levels of stimulus completeness (mean); right, bias measure (see 2.2.4 for a detailed explanation) - difference in accuracy scores for learned minus new stimuli calculated separately for each participant (mean); positive values indicate a bias toward pattern completion, significant differences from 0 are indicated with \*.

Secondly, assessing pattern separation, for the MST, B.L. did not perform differently from controls in identifying targets ( $t_{(19)} = -0.24$ ,  $p = 0.82$ ), or foils ( $t_{(19)} = -0.98$ ,  $p = 0.34$ ), but was considerably less accurate in identifying lures ( $t_{(19)} = -2.50$ ,  $p = 0.02$ ). That is, he thought a lure was old almost five times more often than correctly recognizing that it was similar (see Figure 15). Furthermore, B.L.'s Lure Discrimination Index (LDI) – a measure thought to resemble pattern separation ability (see Methods 4.2.2 for details) – was almost zero, distinguishing him from the controls ( $t_{(19)} = -2.18$ ,  $p = 0.04$ ). When comparing the two tasks, the MIC's bias score and the MST's LDI were contrasted and tested with the Revised Standardized Difference Test (Crawford and Garthwaite 2005). It revealed opposing effects on both tasks between B.L. and controls (see Figure 16). Taken together, B.L. showed a strong bias towards pattern completion, along with an impairment in pattern separation.



**Figure 15. Ch. 4 - Performance on the Mnemonic similarity task (MST).** Accuracy (mean  $\pm$  SE) for learned images (targets), new images that are similar to targets (lures), and new images unrelated to targets (foils).



**Figure 16. Ch. 4 - Bias scores for the Mnemonic similarity task (MST) and the Memory Image Completion task (MIC).** For the MST, data are represented as mean LDI scores;  $\pm$  SE for controls. For the MIC, data are represented as mean bias scores for the degraded or masked images;  $\pm$  SE for controls. Significant differences between controls and BL are indicated with \* $p < 0.05$ , \*\* $p < 0.01$ , and \*\*\* $p < 0.001$ .

#### 4.4. DISCUSSION

In summary, B.L., a patient with selective bilateral DG lesions demonstrated a specific impairment in discrimination abilities for similar stimuli alongside a response tendency in favour of familiar stimuli. Furthermore, B.L.'s ability to recognize previously learned items/targets in both the MST and the MIC, was similar to controls. Taking into account the specificity of his behavioural performance, and considering computational theories (O'Reilly and McClelland 1994; Treves and Rolls 1994) and findings in rodents (Neunuebel and Knierim 2014) this can be interpreted as a pattern separation deficit paired with a bias towards pattern completion. The latter suggests that B.L.'s CA3 is likely intact, although this cannot be proven with absolute certainty on the MR scans. In this light however, reference should be made to recent findings that population activity in proximal CA3 (the part of CA3 which is closest to DG) demonstrates computational pattern separation, whereas distal CA3 demonstrates computational pattern completion (Lee et al. 2015). Thus, if B.L.'s CA3 is impacted, it is likely the part that is physically and functionally closer to DG, i.e. that executes pattern separation.

This study is one of the first to lend direct support to the theories behind the two behavioural tasks, i.e. that the MST and MIC, respectively approximate pattern separation and completion. To be more precise, B.L. clearly presented with a shortfall to dissociate very similar stimuli, that is, to form non-overlapping representations of said items. Simultaneously, he did not have difficulties retrieving learned items from partial cues, in other words, completing the input to the formed representation, but he also tended to do that for stimuli he had not seen before. For these reasons, we assume that B.L.'s DG does not process and project information onto his CA3, which in turn overly reactivates existing memory traces. These findings are in line with computational models (Treves and Rolls 1994; O'Reilly and McClelland 1994), and rodent lesion and genetic studies (Neunuebel and Knierim 2014; McHugh et al. 2007; Gilbert et al. 2001; Nakazawa et al. 2002), which have identified the DG as necessary for pattern separation and CA3 for pattern completion. However, it should be noted here that some computational models and fMRI findings attribute pattern separation roles to CA1 (Rolls 2016) or CA3 (Myers and Scharfman 2009), as well as pattern completion to DG (Nakashiba et al. 2012) or CA1 (Lacy et al. 2011; Bakker et al. 2008). Nevertheless, I would like to emphasize that this is the first study to show specific behavioural effects of DG lesions in humans, enabling us to disentangle pattern separation and completion, and proving the validity of simple behavioural tasks to assess these memory processes.

#### 4.5. CONTRIBUTIONS

The study was a cooperation with York University, Rotman Research Institute and the University of Toronto, namely with Stevenson Baker, R. Shayna Rosenbaum, Asaf Gilboa, Fuqiang Gao and Sandra E. Black. It was designed, carried out, analysed and written into a published manuscript by the Toronto groups. Thomas Wolbers discussed the findings, and edited the paper before publication. I provided the MIC task, helped analyse it, discussed the findings, edited the paper before publication, and rewrote it to fit the means of this thesis.

# 5. THE MIC IN 7T NEUROIMAGING

## Assessing the neural mechanisms contributing to MIC performance with 7T-fMRI

### 5.1. INTRODUCTION

The following experiment was designed to identify neural activity and specific contributions of hippocampal subfields to solving the task. The previous chapters have provided a behavioural basis (chapters 2 and 3) and some first hints to hippocampal involvement (in chapter 4) in what is interpreted to be pattern completion. Here, I wanted to address this more systematically using ultra-high resolution imaging at 7T.

To reiterate, hippocampal subfields have been suggested to differently contribute to memory processes, that is, in the most prominent view DG performs pattern separation, and CA3 performs pattern completion (O'Reilly and McClelland 1994). This is based on the neuroanatomy and neural behaviour of these regions; in essence, because DG shows sparse firing it is well suited to orthogonalize representations, and because CA3 is heavily interconnected it can increase representational overlap through auto-association (see Introduction for detail 1.3). This theory has also received direct evidence from the rodent literature (Neunuebel and Knierim 2014).

In human research, however, the matter is more complicated. Owing to technical limitations, DG and CA3 could not be differentiated so far, with the exception of Bonnici et al. (2012a, 2012b), although their CA3 segmentation largely consists of the fimbria which is anatomically not correct (see chapter 6.2.5.2). Most studies, yet, have reported an involvement of the DG/CA3 complex in pattern separation, and CA1 in pattern completion by virtue of repetition suppression again using the MST (for review, see Yassa and Stark 2011).

Using 7T fMRI, it is possible to achieve higher resolutions ultimately providing the possibility to differentiate even smaller brain regions like CA3 and DG functionally. Employing 7T fMRI, a more recent study from Magdeburg has carved out the specific involvement of DG in pattern separation (Berron et al. 2016). Likewise, the aim of this study was to identify hippocampal subfields involved in pattern completion. Theoretically, as discussed above, CA3 should stand out as a major contributor. More specifically, assessing memory processes with the MIC, it is assumed that when a partial image needs to be identified, pattern completion should occur recruiting CA3. Additionally, the age-related behavioural bias towards pattern completion, is assumed to be concurrent with CA3 hyperactivity (in line with Wilson et al. 2006; Yassa et al. 2011a). In order to test these hypotheses, the MIC was slightly adapted to fit neuroimaging requirements and was tailored for multivariate analyses by increasing trial numbers, reducing conditions, jittering timing and balancing task difficulty (see Methods 5.2.2 for details).

Multivariate pattern analyses were used because they can capture stimulus-specific representations and are less sensitive to variability in activity levels between subjects (for detailed benefits of MVPA, see Davis et al. 2014). Crucially, univariate analyses cannot pick up fine-grained differences in activity patterns within a given ROI, but rather identify differences in the mean population responses over all voxels. Given that pattern separation and completion are assumed to orthogonalize or increase overlap of specific representational patterns, multivariate analyses are far better suited to identify differences between whole activity patterns in contrast to average differences assessed by univariate analyses. Here, I employed multi-voxel pattern similarity analysis (adapted from Kriegeskorte et al. 2008), to see how stimulus-specific brain activity would be correlated in each hippocampal subfield. Hypothesizing that a given stimulus has a corresponding representation as picked up by fMRI, correlations of activity for trials showing the same stimulus should be high as opposed to correlations with activity for other stimuli. Similarly, assuming that a partial image is completed towards its full representation, the corresponding activity pattern should correlate strongly with the activity elicited by the original full stimulus. Employing this method, and based on the neuroanatomical literature, I assumed that CA3 would present with high correlations for matching trials, and low correlations for non-matching trials. Moreover, DG - if serving pattern separation - should show lower correlations altogether. CA1 and subiculum should exhibit patterns associated with the final response assuming that these regions integrate the processed information coming in from CA3 and DG (in line with Lee et al. 2004). Additionally, I wanted to investigate the underlying activity differences associated with age. Given that older adults have a tendency to choose familiar responses over new ones, it is assumed that they have a bias toward pattern completion (Wilson et al. 2006), which should manifest in heightened correlations for activity of new stimuli when they are falsely identified as old (see Methods on multivariate analyses 5.2.5.2 for a detailed description of all investigated correlations). Critically, I wanted to test one major assumption in theories about cognitive aging stating that CA3 should be hyperactive during mnemonic processing (Wilson et al. 2006).

## 5.2. METHODS

### 5.2.1. SUBJECTS

All participants were recruited by the DZNE Bonn, and pre-screened for MR-compatibility on the phone. When eligible, participants received a detailed study description, consent form and health questionnaire by mail, which they were asked to read and fill in before coming in for testing. On the test day, participants were carefully briefed about the experimental procedures, and had the opportunity to resolve any remaining questions about the study before signing the consent form. Their cognitive abilities were then assessed using the MoCA (Nasreddine et al. 2005), and they were excluded from further participation if their score was below 23 (Luis et al. 2009). Finally, participants were again screened for MR-compatibility, and respectively cleared for MRI testing by a study doctor. Overall, 47 participants underwent the scanning

session. Each brain was checked by a study doctor for structural abnormalities (none were observed). Five participants were excluded due to preliminary abortion of the experiment caused by back pain, ear ringing, dizziness, excessive motion or scanner problems, resulting in a final study population of 21 young (20-34 years old; 12 females) and 21 older adults (60-75 years old; 10 females). The study received approval from the Ethics Committee of the University Clinic of Bonn. All participants received monetary compensation of 10€/h.

**5.2.2. MATERIALS**

This experiment was an extension of the MIC (see chapters 2.2, 3.2.2 and 4.2.2), adjusted to fit requirements of neuroimaging analyses (e.g. less stimuli, more repetitions, jittered timing, etc.). Here, I used 12 line-drawn scenes from the previous set, 4 of which had to be learned, 4 were used as foils to test learning, and 4 were used as novel images. Of the 8 stimuli used in the test phase (i.e. 4 learned and 4 novel stimuli), I created 2 masked versions each (21% and 12% visible of the original image; see Table 4). I chose those masking levels, because they produced the strongest effects in the behavioural version of this paradigm (see chapter 2.3.2). Each stimulus was repeated 12 times (4 times within each of 3 sessions, see Procedure 5.2.3). Contrary to all previous experiments with this paradigm, the complete versions of the novel stimuli were never shown. This was done to prevent ceiling effects, as the more frequent repetitions and less diverse masking of each stimulus decreased the overall difficulty of the task.

**Table 4. Ch. 5 - stimulus material for the test phase.**

Stimuli	Full (48 trials)	Partial (192 trials)
<b>Learned</b>	4 scenes (kitchen, bedroom, library, dining room) 12 repetitions each	4 scenes masking level 21% 12 repetitions each
		4 scenes masking level 12% 12 repetitions each
<b>New</b>	none (office, living room, classroom, restaurant)	4 scenes masking level 21% 12 repetitions each
		4 scenes masking level 12% 12 repetitions each

**5.2.3. PROCEDURE**

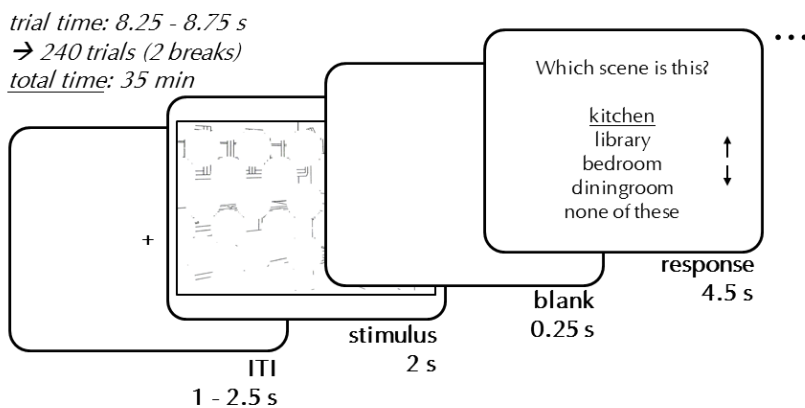
Prior to scanning, participants learned 4 scene exemplars, each of which was presented 3 times on a 17" computer screen for 2 seconds within a random sequence with the other scene exemplars. Preceding each image, a verbal label of the scene was shown for 1 second. Based on a recognition criterion, it was ensured that participants had learned the items, i.e. they had to recognize and correctly identify each exemplar 3



times when presented intermixed with 4 foil images (see also chapter 2.2.3 for more details).

All participants received a short training session outside and inside the scanner to familiarize them with the task described below. They also saw a short reminder of each learned item before scanning started.

Inside the scanner, participants had to perform a recognition memory task (see Figure 17), which was shown on an MR-compatible screen behind the scanner that participants could see via a mirror inside the head coil. In each trial, a jittered fixation cross appeared for a random duration between 1 and 2.5 seconds. After that, an image was presented on a grey background for 2 seconds, followed by a short blank of 0.25 seconds, after which the response screen appeared for 4.5 seconds. Participants had to indicate which scene they had just seen by moving a horizontal line cursor to the appropriate response option (see Figure 17). By default, the cursor was positioned at the question, to not influence the participants' choice. For the same reason, response options were always shown in a random order. Participants could move the cursor up and down by button presses of their right index and middle finger. They had to register their final choice via button press with their left index finger. Once they had done so, the cursor changed colour from white to black. After responses were made, the response screen stayed on until the 4.5 seconds were over, before the next trial started. The image in each trial could be learned or novel, and was presented in either masked (21% or 12% visible) or complete form (learned items only; see Table 4). Participants performed 3 sessions of this task, each of which consisted of 80 trials and lasted approximately 12 minutes.



**Figure 17. Ch. 5 - experimental design of the test phase.** Each trial started with a jittered fixation cross of 1-2.5 s. Then the stimulus was presented for 2 s each, followed by a 0.25 s blank. On the response screen which was present for 4.5 s, a white line bar was visible underneath the question. Participants could move that line up and down with buttons in their right hand, once they wanted the final answer to register, they pressed a button in their left hand and the line turned black. The response screen stayed on until the 4.5 s were over.

#### 5.2.4. MRI ACQUISITION

Imaging data were collected at the German Centre for Neurodegenerative Diseases (DZNE) in Bonn on a 7 Tesla MR scanner (Siemens, Erlangen, Germany) with a 32-channel head coil (Nova Medical, Wilmington, MA, USA). A 1 mm isotropic whole-brain structural volume (MPSAGE; TE = 2.73 ms, TR = 2500 ms, TI = 1100 ms, flip angle = 7°; see Stöcker and Shah 2006; Brenner et al. 2014) was acquired for slab positioning of the following scans. Afterwards, every participant underwent three functional scanning sessions, each of which consisted of approximately 330 volumes (~12 min) depending on the length of the paradigm (variable due to randomized ISIs,

see Procedure 5.2.3). A T2\*-weighted 2D echo planar image (EPI) slab of 0.8 mm isotropic resolution was positioned parallel to the long axis of the hippocampus, and acquired in an odd-even interleaved fashion (28 slices, TE = 22 ms, TR = 2000 ms, flip angle = 85°, FOV = 205 mm, partial Fourier = 5/8, parallel imaging with grappa factor 4, bandwidth = 1028 Hz/Px, echo spacing = 1.1 ms). EPIs were motion and distortion corrected online via point spread function mapping (see In and Speck 2012). To facilitate coregistration of functional and structural images, one whole brain EPI of 1.6 mm isotropic resolution was acquired (TE = 22 ms, TR = 5000 ms, flip angle = 90°, FOV = 205 mm, partial Fourier = 5/8, parallel imaging with grappa factor 4, bandwidth = 2056 Hz/Px, echo spacing = 0.65 ms).

For structural segmentation, a partial turbo spin echo (TSE) T2-weighted volume was acquired oriented orthogonal to the long axis of the hippocampus (resolution = 0.4 × 0.4 mm, 55 slices, slice thickness = 1 mm, distance factor = 10%, TE = 76 ms, TR = 8020 ms, flip angle = 60°, FOV = 224 mm, bandwidth = 155 Hz/Px, echo spacing = 15.2 ms, TSE factor = 9, echo trains per slice = 57).

---

### 5.2.5. MRI ANALYSIS

FMRI pre-processing and statistical modelling were performed with Statistical Parametric Mapping (SPM8; Wellcome Department of Cognitive Neuroscience, University College, London, UK). The first and last trial of each session needed to be removed because of elevated noise in the MR signal resulting in 234 total trials which were used for analyses. All EPI volumes were already unwarped and aligned across all three sessions through online distortion and motion correction via point spread function mapping (In and Speck 2012). The motion parameters were extracted from the image headers with an in-house MATLAB script (written by Matthias Stangl).

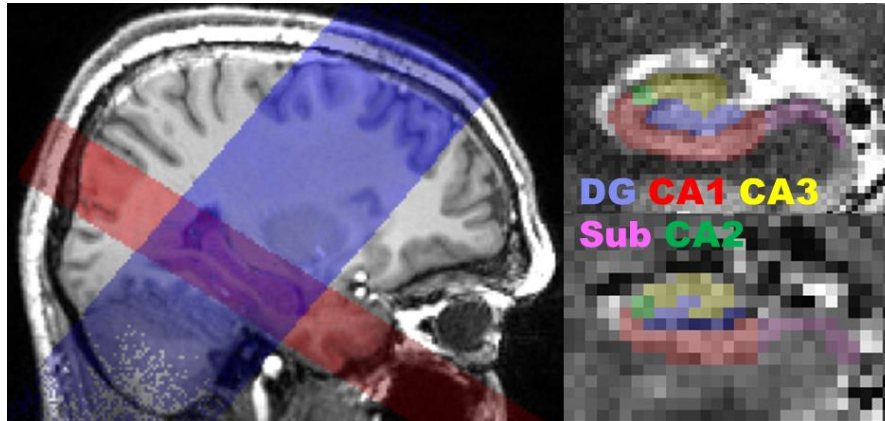
The high-resolution structural T2 slab was used for manual segmentation (see next paragraph 5.2.5.1), and therefore needed to be coregistered to the functional runs which is computationally more demanding than a standard coregistration as both sequences are only partial volumes with very small overlap (see left image in Figure 18). To that effect, I first calculated a mean image over all EPI slabs. The whole-brain EPI was then re-oriented to match the orientation of the mean EPI (using coregister: estimate only). Likewise, the T2 was then coregistered to the re-oriented whole-brain EPI (coregister: estimate only). The orientation information in the header of the T2 from this point was later copied to the segmented ROIs. Finally, the T2 was resliced to fit the mean EPI (coregister: reslice only) for further analyses.

---

#### 5.2.5.1. HIPPOCAMPAL SEGMENTATION

First, all hippocampi with subfields CA1, CA2, CA3, DG and subiculum (Sub) were automatically segmented on the original T2-weighted volumes using Automatic Segmentation of Hippocampal Subfields (Yushkevich et al. 2015b). Afterwards, each segmentation was manually adjusted according to the protocol of Wisse et al. (2012) using itk-SNAP ([www.itksnap.org](http://www.itksnap.org); Yushkevich et al. 2006). However, additionally, the endfolial pathway was used as a boundary between CA3 and DG, because it is true to

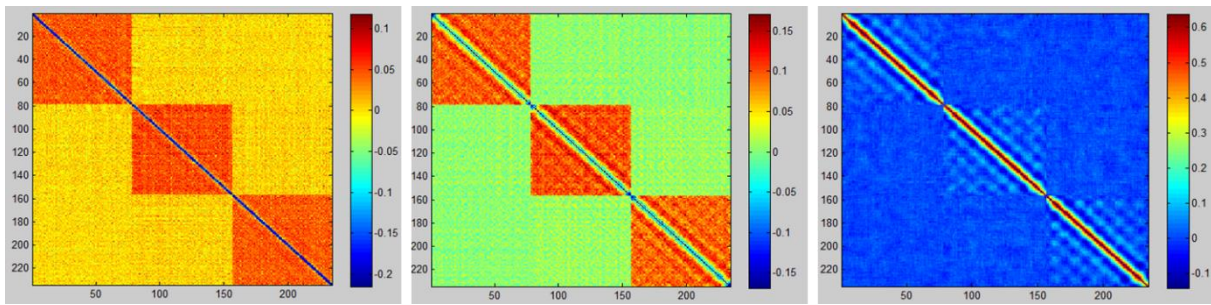
actual neuroanatomy and was clearly visible on all our scans (see also 6.2.5.1). After segmentation, the masks of CA2 and CA3 were combined for further analyses, because CA2 is so small that it only spans 2-4 voxels per slice in functional space (exemplary, see right lower image in Figure 18).



**Figure 18. Ch. 5 – MRI volumes and segmentations.** (Left) A 0.8 mm isotropic EPI-slab aligned to the hippocampal main axis (shaded red area), and a  $0.4 \times 0.4 \times 1.0 \text{ mm}^3$  T2-slab orthogonal to the hippocampus (shaded blue area) are overlaid on a skullstripped T1. (Right) Hippocampal segmentation of CA1, CA2, CA3, DG and Sub on the original T2 (top) were coregistered to the mean functional EPI (bottom). Note, that for analyses CA2 and CA3 were combined to one mask.

#### 5.2.5.2. MULTIVARIATE ANALYSES

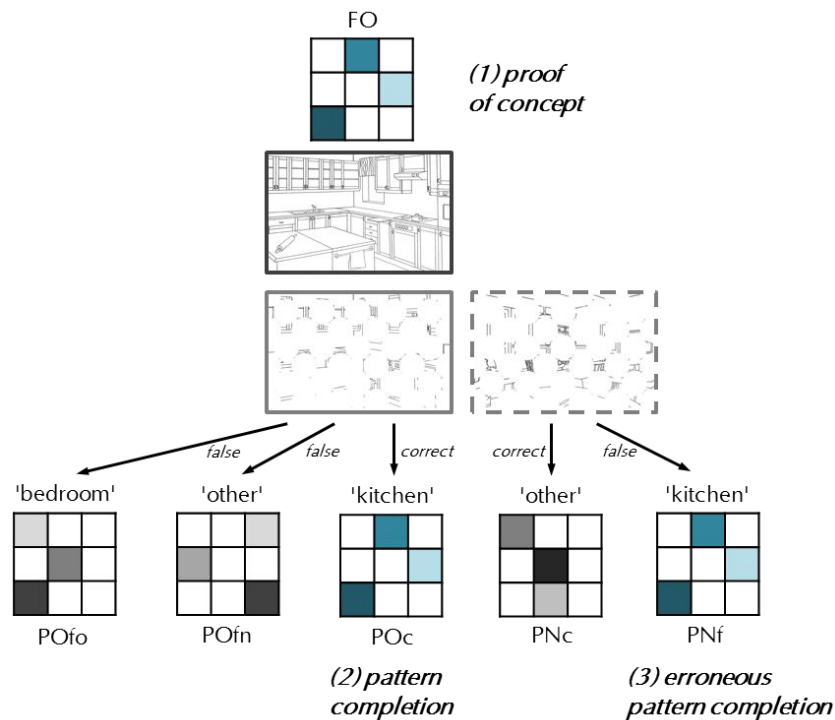
Unsmoothed data were used for multivariate analyses to keep the specificity of the subfield responses gained by our high resolution. Single trial models were implemented following Mumford et al. (2012) because these types of models reduce temporal overlap between trials. That is, a separate GLM was estimated for each trial resulting in 234 GLMs. For each GLM, stimulus onset of the respective trial was modelled as the main regressor of interest; additionally, one regressor with the corresponding response window, one regressor with all other trials, one with all other response windows, and six movement parameters were included. This was done session-specific. Duration of each regressor was set to 0. The use of these particular GLMs were the result of extensive testing confirming that they were the best choice to eliminate temporal overlap (in line with Mumford et al. 2012) even though they considerably decreased the maximum correlation values (see left plot in Figure 19 for corresponding correlations). Prior models were dismissed for various reasons: e.g. (1) 1 GLM with 234 separate regressors resulted in 'ringing' across the correlations of the estimated regressors first thought to reflect an external artefact but which is a side-effect of high-pass filtering in single trial models (see right plot in Figure 19); (2) 234 GLMs without extra response modelling resulted in a strong decorrelation of the regressors closest in time and again slight 'ringing' (see middle plot of Figure 19).



**Figure 19. Ch. 5 – Single trial correlation matrices dependent on model specifications.** Values denote Fisher- $z$  transformed  $r$ -values from Pearson correlations, exemplary in subiculum. (Left) Model finally selected for this analysis: 234 GLMs with each trial modelled separately and the corresponding response window as another regressor resulting in the lowest maximal correlations but fewest temporal artefacts; (Middle) 234 GLMs with each trial modelled separately resulting in slight 'ringing' (repetitive parallel lines to diagonal) and strong decorrelation of temporally close trials (green areas next to diagonal); (Right) 1 GLM with 234 separate trial regressors resulting in strong temporal correlations and 'ringing'. Each visible square (mainly red in the two left images) denotes a session.

For multi-voxel pattern similarity analysis (Kriegeskorte et al. 2008), the individual parameter estimate (beta) of each trial was extracted for all ROIs separately for both hemispheres, i.e. left and right hippocampal subfields: CA1, CA2/3, DG, subiculum. Afterwards, all trial-by-trial Spearman correlations (as suggested by Kriegeskorte et al. 2008) were calculated per ROI, resulting in a  $234 \times 234$  correlation matrix for each ROI. R-values were then Fisher- $z$  transformed and averaged across trial pairs of interest. In contrast to the original paper, I used the pure correlations as a measure of similarity rather than dissimilarity (1-correlation; Kriegeskorte et al. 2008). It should be noted that by now the Mahalanobis distance has been suggested as a better similarity measure (Walther et al. 2015), however re-analyses of the current data using Pearson correlations, Euclidean and Mahalanobis distances did all yield similar results.

Because within-session correlations are susceptible to increased false positives (Mumford et al. 2014), I subsequently only used between-session correlations (trials from session 1 correlated with trials from session 2 and 3, etc.). Furthermore, I matched trial numbers across condition comparisons to exclude the possibility that elevated correlations could be explained by more data contributing to the correlation as compared to another condition. That is, if one condition had less trial pairs than another condition, a random subset of trials was selected individually per participant. Further on, it was assumed that the representation of a particular image was present in the corresponding trial's activity pattern across all voxels in a given ROI. More specifically, I selected all trials of complete stimuli as the respective default patterns (activity elicited by full presentations of the kitchen, library, bedroom or dining room) to correlate with other conditions as follows. For an illustration of the analysis design see Figure 20 (including all the following comparisons).



**Figure 20. Ch. 5 – Response-based analysis design.** Checkered squares illustrate simplified activity patterns in a given set of voxels corresponding to a specific stimulus. The representation of a complete stimulus (FO = full old) is shown at the top (in blue). Recognition of the same stimulus is assumed to elicit similar activity patterns through pattern completion, either when a partial version of the original cue is correctly identified (POc = partial old correct), or another stimulus is falsely assumed to show the original cue (PNf = partial new false; both also indicated by blue patterns). In all other cases, different representations should be detected. Analyses are based around comparisons of stimuli represented by the blue patterns with all other options. In (1) *proof of concept*, all within-FO correlations are compared to all between-FO correlations. In (2) *pattern completion*, within-POc correlations are compared to between-POc, between-PNc and within-POfn correlations. In (3) *erroneous pattern completion*, within-PNf correlations are compared to between-POc, between-PNc and within-POfn correlations. For more detailed descriptions, see text.

First, I wanted to provide a *proof of concept* that representations of specific stimuli could be detected using single trial models. To achieve that, I only used trial pairs of full stimuli to compare within-stimulus to between-stimulus correlations, e.g. within: every kitchen trial was correlated with all other kitchen trials, and between: every kitchen trial was correlated with a subset of all trials showing library, bedroom or dining room. The idea here is that the representation of a kitchen should be more similar to the representation of other kitchen trials than to all other stimuli, resulting in higher correlations.

Second, if *pattern completion* occurred, a partial image should reinstate (complete) the representation of the corresponding full stimulus. Theoretically, this should be the case when partial images that had been learned previously, are correctly identified. Therefore, I used correlations of the full learned images with their corresponding correctly identified learned partial images (within-POc; e.g. kitchen with kitchen) and compared them to three different conditions where pattern completion should not have occurred or failed: (1) correctly identified learned partial images that did not correspond to the full images (between-POc; e.g. kitchen with bedroom); (2) correctly identified new partial images (between-PNc; e.g. kitchen with new stimulus restaurant); (3) partial learned images that were incorrectly judged as new (misses, within-POfn;

e.g. kitchen with kitchen identified as restaurant). Partial trials included both the 21% and 12% masked stimuli to increase trial numbers.

Third, sometimes *erroneous pattern completion* can occur. This should be particularly frequent in older adults. For example, when new images are falsely identified as one of the learned ones, it is assumed that pattern completion should also play a role even though leading to false identification (false alarms). Therefore, all correlations of falsely identified new images with the corresponding full images (between-PNf; kitchen with new stimulus restaurant identified as kitchen) were contrasted to the same three conditions as above (1-3) where pattern completion should be absent.

All comparisons were tested with paired t-tests on the Fisher-z transformed  $r$ -values resulting from the correlations, separately for each ROI.

---

### 5.2.5.3. UNIVARIATE ANALYSES

Univariate analyses were done in addition after multivariate results had been obtained. They can inform about more global processes and regions contributing to certain aspects of the task without having to rely on specific activity patterns and with less dependency on behaviour. Critically, age differences can be assessed by looking at contributions of certain brain regions in general but also with regard to the processing of different conditions.

Functional images were smoothed with a 1.6 mm full-width-half-maximum Gaussian kernel. To prevent obliterating the data through multiple analysis steps, all computations were done in single subject space and normalization to group-level on the individual contrast-images was only done afterwards (see below). General linear models (GLM) were calculated in native space, including all three sessions as separate runs.

#### *Assessing performance-dependent processes possibly engaging pattern completion*

First, in the wake of the original design, a GLM was set up accounting for behavioural differences and to tackle recognition memory that likely engaged pattern completion. Here, I calculated five condition regressors split according to stimulus type, completeness and performance, i.e. old and new, full and partial, correct and false trials. Partial regressors each comprised the 21% and 12% masked stimuli together to increase trial numbers. Because all participants were almost 100% accurate for the full stimuli, I did not model false trials for full stimuli. This resulted in five condition regressors per session: *correct full old (FOc)*, *correct partial old (POc)*, *correct partial new (PNc)*, *false partial new (PNf)* and *false partial old split up into old/new errors (POfo; POfn)*. Seven more regressors were included per session: one regressor modelling all response windows (the time window immediately following image presentation where participants could make a response), and six movement parameters. This resulted in a total of 36 regressors (12 regressors per session). All regressors were modelled at the respective stimulus onsets convolved with the hemodynamic response function (HRF) and a duration of 0 s. Two elderly participants did not respond correctly

to any new masked trials in any of the sessions, thus, their data did not contribute to the respective regressors and follow-up contrasts.

#### *Estimating pattern completion as an active process*

First, I contrasted conditions where I assumed pattern completion would be involved with conditions where it should not be. Consequently, this is the case for learned partial images that are correctly identified, but also for new partial images that are mistaken for learned images, and learned images that are incorrectly thought to depict other learned stimuli. These conditions were separately contrasted with correct identifications of new stimuli where I did not expect pattern completion to be involved (*correct partial old > correct partial new, POc > PNC* and *false partial new > correct partial new, PNf > PNC*; *false partial old identified as old > correct partial new; POfo > PNC*). These contrasts assume an active process of pattern completion, that is, there should be more activity when it occurs. Additionally, this contrast could inform about potential age differences linked to CA3-hyperactivity.

#### *Estimating pattern completion by repetition suppression*

Alternatively, employing the concept of repetition suppression, I calculated the opposite three contrasts also trying to get at pattern completion processes (*correct partial new > correct partial old, PNC > POc* and *correct partial new > false partial new, PNC > PNf* and *correct partial new > false partial old identified as old, PNC > POfo*). The reasoning here is as follows: when a stimulus is repeated, activity levels usually drop due to adaptation (Krekelberg et al. 2006). Thus, if a partial stimulus is recognized as one of the learned stimuli, i.e. pattern completion has occurred, there should be less activity than for a stimulus judged to be new (see also explanation in the Introduction 1.3.3).

#### *Performance-invariant processes involved in the MIC*

Further on, a second GLM was calculated to make more general assertions on neural activity associated with the MIC. Per session, three condition regressors were modelled irrespective of behaviour, i.e. *full old, partial old, partial new*. The same additional regressors were included as for the first model (response windows, movement parameters). Following that, I calculated three contrasts estimating more general mechanisms involved in recognition memory: stimulus completeness/visibility, novelty, and unspecific retrieval. The effect of visibility should become apparent in comparing activity levels for full stimuli with partial ones. To exclude any other concomitant signal change relating to stimulus type, I only contrasted learned images with each other (*full old > partial old, FO > PO*). Similarly, to assess stimulus novelty, I used only the partial images to see which regions would be more engaged by new images as opposed to learned ones (*partial new > partial old, PN > PO*). Lastly, in a more exploratory attempt to identify regions that would be more activated by learned stimuli as opposed to new ones, I calculated the last contrast again excluding added effects of visibility (*partial old > partial new, PO > PN*).

### Group-level analysis

For effect localization and visualization on the group level (normalization), a study-specific population template was created combining all T2-weighted scans using nonlinear diffeomorphic mapping with Advanced Normalization Tools (ANTs; command line script `buildtemplateparallel.sh`; Avants et al. 2011). One older participant was excluded from the template due to a warping failure of the software (i.e. the brain was too different to be mapped onto the group template). Afterwards, individual mean functional images were registered and resliced along with the contrast images to match the individual T2 images with SPM8. Subsequently, I warped the individual resliced contrast images into the template space with ANTs using the transformation matrices derived from the template building process. I could then use the aligned contrast images for second-level group analyses in SPM8. As a first step, I used a flexible factorial design with factors age group (young, old) and condition (FO/PO/PN and respectively FOc/POc/PNc/POfo/POfn/PNf) to identify main effects and interactions, and subsequently used one-sample and two-sample t-tests to test more specific differences.

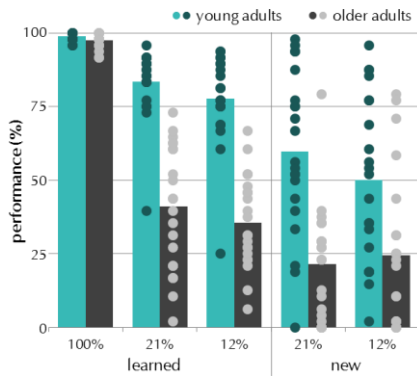
## 5.3. RESULTS

### 5.3.1. BEHAVIOURAL RESULTS

Recognition ability was assessed by looking at accuracy scores separately for learned and new items at different levels of stimulus completeness, respectively (see Figure 21). Because there were no complete new stimuli, I first ran a three-way mixed ANOVA on the incomplete stimuli only, with a between-subjects factor of age (young, old), and two within-subject factors (stimulus completeness: 21%, 12%; stimulus type: learned, new). Young adults performed better than older adults (main effect of age:  $F_{(1,38)} = 92.851$ ,  $p < 0.001$ ), and both groups had more difficulties identifying new stimuli when compared to learned stimuli (main effect of stimulus type:  $F_{(1,38)} = 31.177$ ,  $p < 0.001$ ), in contrast to all previous versions of the study (see 2.3.1, 3.3.1.1, and 4.3). It is of note that overall variance on the task was also considerably higher to all previous studies especially in young adults and for new stimuli (see Figure 21). Additionally, stimulus completeness modulated accuracy (main effect of stimulus completeness:  $F_{(1,38)} = 12.705$ ,  $p = 0.001$ ), and differently affected the groups' responses as revealed by a two-way interaction (age  $\times$  stimulus completeness:  $F_{(1,38)} = 6.860$ ,  $p = 0.013$ ); i.e. reduced stimulus completeness resulted in lower accuracy, except older adults slightly improved from 21% to 12% new stimuli (see Figure 21). Not surprisingly, I found a difference in accuracy between full and partial images, as revealed by another ANOVA on the learned stimuli only (main effect of stimulus completeness:  $F_{(2,76)} = 250.093$ ,  $p < 0.001$ ) which also confirmed the previous effects (main effect of age:  $F_{(1,38)} = 113.416$ ,  $p < 0.001$ ; age  $\times$  stimulus completeness:  $F_{(2,76)} = 78.785$ ,  $p < 0.001$ ). Post-hoc independent t-tests revealed that young and older adults performed equally well for the complete stimuli ( $t_{(29,072)} = 2.000$ ,  $p = 0.055$ ), but differed in performance for both incomplete learned and new stimuli (after Holm-Bonferroni multiple comparisons correction; all p



< 0.005; level 21%:  $t_{learned(38)} = 9.229$ ,  $t_{new(38)} = 5.785$ , level 12%:  $t_{learned(33.553)} = 11.091$ ,  $t_{new(37.448)} = 3.185$ ). Interestingly, older adults again showed a specific tendency to choose a familiar response over a new one. For learned stimuli, older adults' false alarm rate was a lot higher than that of young adults, although they had more misses in total than false alarms (see Table 5). Unfortunately, there were not enough error trials for learned stimuli to conduct a statistical test, wherefore I only present descriptive data.



**Figure 21. Ch. 5 – Performance measures.** Performance for both age groups, separately for learned and new stimuli for the different levels of stimulus completeness (mean).

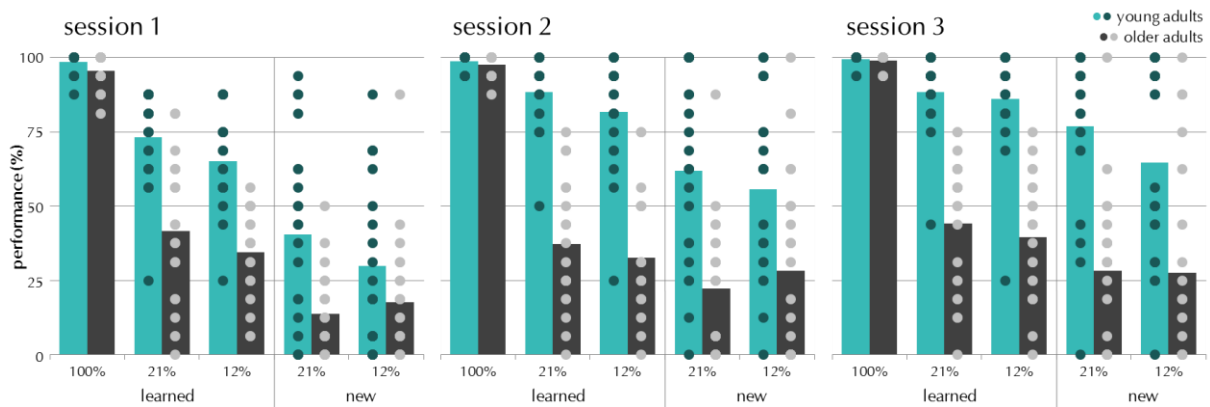
**Table 5. Ch. 5 - false alarm rates for learned stimuli.**

stimulus completeness	false alarms - mean (SE)	
	young adults	older adults
100%	0.03 (0.00)	0.04 (0.00)
21%	0.09 (0.01)	0.40 (0.03)
12%	0.12 (0.02)	0.42 (0.04)

False alarms and misses add up to 1, so that values higher than 0.5 indicate more false alarms, and values lower than 0.5 indicate more misses; values do not comprise the data of all participants since not all of them made errors for each completeness level.

More to the point, older adults hardly identified a new stimulus as such but rather attributed it to an old one. To further investigate this effect, I looked at the distribution of false alarm options for new stimuli. One false response option should have been chosen over all other false responses. Therefore, I calculated the group-average frequencies of all false choice options and sorted them from most (FA 1) to least (FA 4) chosen option for each new stimulus, and averaged across all new stimuli afterwards. A  $\chi^2$ -test of goodness-of-fit on the 4 false alarm options ( $\chi^2 = 218.436$ ,  $df = 3$ ,  $p < 0.001$ ) revealed that older participants indeed chose one particular false response option most often (FA 1) and did not simply guess more. I also checked whether there was one dominant option for each stimulus, by contrasting the most frequent false alarm against the average of the other false alarms per stimulus (stimulus 'office':  $\chi^2 = 22.427$ ,  $df = 1$ ,  $p < 0.001$ ; stimulus 'class room':  $\chi^2 = 19.358$ ,  $df = 1$ ,  $p < 0.001$ ; stimulus 'restaurant':  $\chi^2 = 22.73$ ,  $df = 1$ ,  $p < 0.001$ ; stimulus 'living room':  $\chi^2 = 5.721$ ,  $df = 1$ ,  $p = 0.017$ ). This is evidence that older participants completed towards the stimulus perceived as most similar, rather than simply guessed more.

Furthermore, I checked whether there were any learning effects over the course of the experiment. Since the test phase was split up into three sessions, I objected the data to a four-way mixed ANOVA with the same factors as above, and an additional factor of session. The plots (Figure 22) show that older adults improve less across sessions than young adults do, which was confirmed by a main effect of session ( $F_{(2,76)} = 56.258$ ,  $p < 0.001$ ), and a two-way interaction (age  $\times$  session:  $F_{(2,76)} = 21.666$ ,  $p < 0.001$ ). Post-hoc independent t-tests showed the same pattern as for the mean across all three sessions (see above).



**Figure 22. Ch. 5 - Performance measures by session.** Performance for both age groups, separately per session, for learned and new stimuli and for the different levels of stimulus completeness (mean).

Due to the responses made by way of cursor movements, I did not analyse reaction times as they were no longer meaningful because response options were randomized and could only be selected successively.

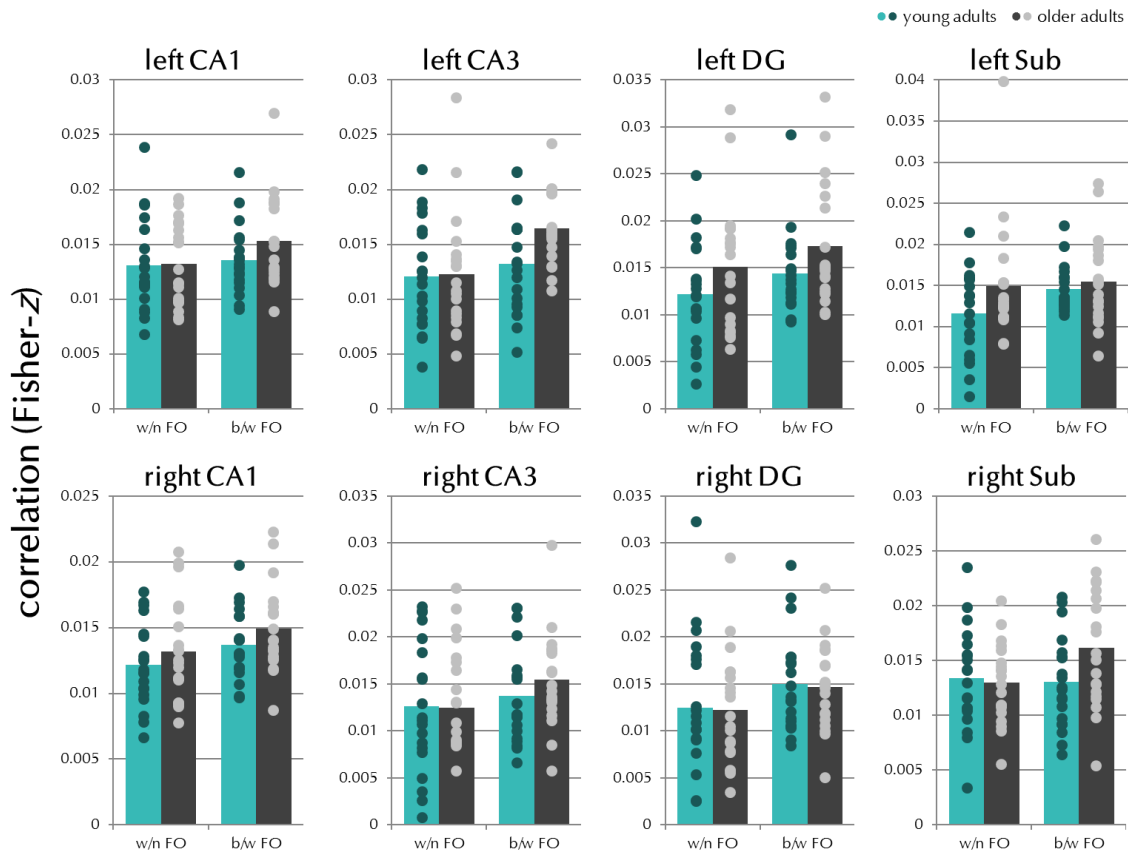
Altogether, these findings show that recognition ability declined with reduced stimulus completeness, it was worse for new stimuli, and its decrease was stronger in older adults. Again, older adults showed a higher tendency to choose familiar stimuli as responses. Moreover, they learned less over the course of the experiment than young adults did. However, young adults were also less accurate for new stimuli as compared to learned ones. Here, I did not calculate bias scores as in the previous experiments, because the stimulus types (learned or new) were not comparable even-handedly. That is, because new stimuli were never shown in their full form, there were fewer trials for this stimulus type, and participants did not have the chance to make a full representation of a new stimulus.

### 5.3.2. NEUROIMAGING RESULTS

One young and one older participant were excluded from the neuroimaging analyses as they were behavioural outliers (over 3 standard deviations away from their age groups mean performance), because both seemed to have an overall response bias towards "new" in all partial conditions (their removal did not change the statistical results of the behavioural analyses).

#### 5.3.2.1. MULTIVARIATE RESULTS

Paired t-tests on Fisher-z transformed  $r$ -values, did not show any differences between within-stimulus and between-stimulus correlations even before multiple comparison corrections (all  $p > 0.05$ ; exemplary, see Figure 23). Thus, a proof of concept to see whether the representations of an identical stimulus were more correlated than other stimuli could not be established. Unfortunately, no hippocampal subfield seemed to differentiate between any correlations suggesting that they may not carry stimulus-specific representations. Similarly, the other two blocks of comparisons trying to identify successful and erroneous pattern completion mechanisms did not yield any unequivocal results nor could age effects be identified with independent t-tests (all  $p > 0.05$ ).



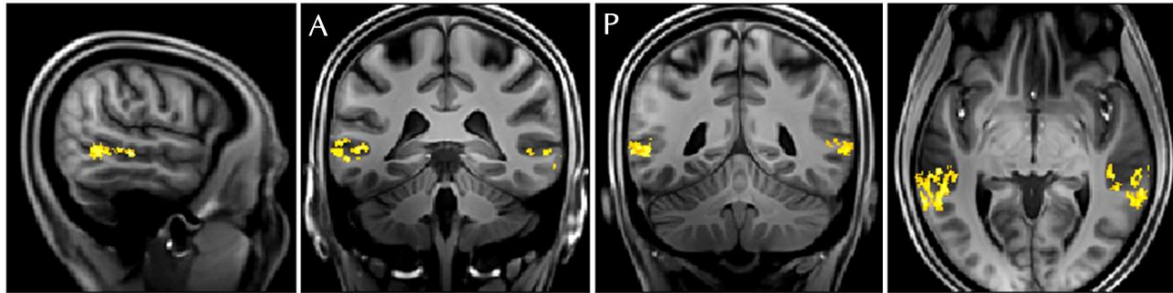
**Figure 23. Ch. 5 – Proof of concept correlations.** Spearman correlations (Fisher- $z$  transformed  $r$ -values) of compared pairs are depicted for each subfield and age-group. Within full old correlations (w/n FO) consist of correlations of e.g. a kitchen stimulus with all other kitchen stimuli. Between full old correlations (b/w FO) consist of correlations of e.g. a kitchen stimulus with a subset of library, bedroom or dining room stimuli (see Multivariate Methods 5.2.5.2 for details). All paired t-tests showed no significant differences ( $p > 0.05$ ).

### 5.3.2.2. UNIVARIATE RESULTS

Because the representational findings were inconclusive, I looked at process contributions of overall brain activity. All contrasts were originally thresholded at  $p_{\text{voxel level}} < 0.001$ , and corrected on the cluster level with family-wise-error (FWE) corrections  $p_{\text{cluster level}} < 0.05$ .

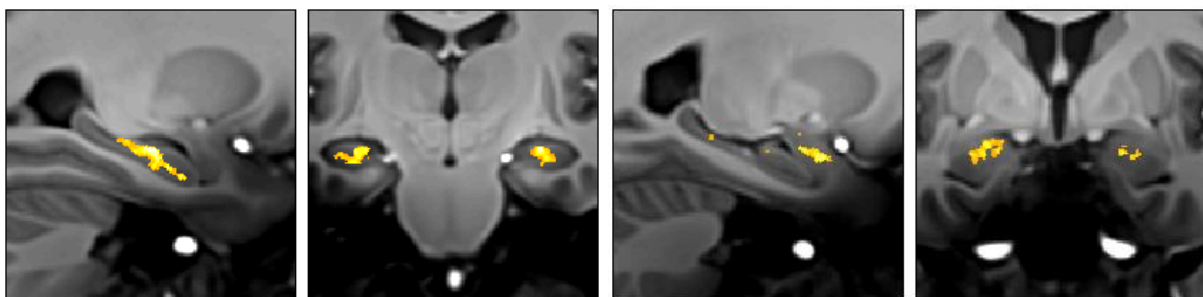
*Pattern completion:* First of all, an active process which would show more activity when pattern completion was needed as opposed to when it was not could not be identified in any region in neither age group with none of the three contrasts. However, repetition suppression indicative of pattern completion was observed in superior temporal sulcus (STS) possibly stretching into lateral occipital complex (LOC) posteriorly – left in young adults and bilaterally in older adults (see Figure 24, and Table 6 and Table 7). The reasoning behind this contrast is: If a stimulus is perceived as having been seen before (pattern completed), adaptation should occur and activity levels should drop. Therefore, correctly identified old stimuli should produce lower activity levels than new stimuli (PNC > POC). Additionally, pattern completion can also be involved in misidentification, i.e. when a new stimulus is perceived as old, or an old stimulus is identified as a different old stimulus, it is assumed that pattern completion is also operating (PNC > PNf and PNC > POfo). Activity associated with these errors was

also observed in the bilateral STS in both age groups. However, activity associated with errors for old stimuli were only observed in older adults (see Table 6 and Table 7), which may be due to their elevated error rate for these particular types of errors (see Table 5). A two-sample t-test confirmed a cluster in left posterior STS which showed higher activity in older adults ( $N_{\text{voxel}} = 643$ ,  $T_{\text{peak}} = 4.44$ ,  $p_{\text{peak}} < 0.001$ ,  $p_{\text{clusterFWE}} < 0.001$ ,  $p_{\text{clusterFDR}} < 0.001$ ).



**Figure 24. Ch. 5 – Pattern completion effects.** The contrast is based on repetition suppression. Greater activity for correctly identified new stimuli than correctly identified old stimuli (PNc > POc) was observed in superior temporal sulcus (A – more anterior slice) possibly stretching into lateral occipital complex (P – more posterior slice), left in young adults, and bilaterally in older adults (thresholded at  $p_{\text{voxel level}} < 0.001$ ;  $k > 100$  voxels; FWE-corrected  $p_{\text{cluster level}} < 0.05$ ).

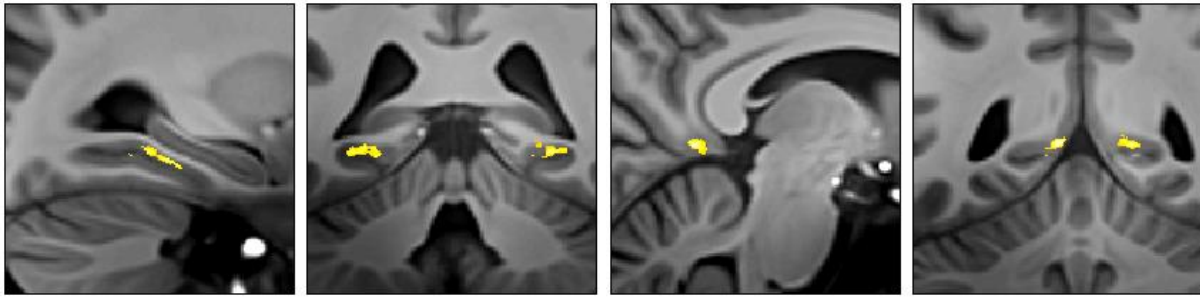
*Retrieval:* Activity associated with a more general retrieval mechanism that differentiated between learned and new stimuli was observed bilaterally in the hippocampus and amygdala of young adults (see Figure 25). These regions exhibited greater activity for old than for new stimuli (PO > PN; see Table 6). The clusters in the hippocampus were fairly big (966 voxels on the left, and 2168 voxels on the right) and included portions of all subfields not rendering it useful to differentiate between them. Interestingly, this contrast did not show any involved brain regions in older adults. A two-sample t-test confirmed a cluster in left hippocampus which showed higher activity in young compared to older adults ( $N_{\text{voxel}} = 216$ ,  $T_{\text{peak}} = 7.092$ ,  $p_{\text{peak}} < 0.001$ ,  $p_{\text{clusterFWE}} = 0.053$ ,  $p_{\text{clusterFDR}} = 0.025$ ).



**Figure 25. Ch. 5 – Retrieval effects in young adults only.** Greater activity for old than new stimuli (PO > PN) was observed in bilateral hippocampus (left) and amygdala (right) in young adults only (thresholded at  $p_{\text{voxel level}} < 0.001$ ;  $k > 100$  voxels; FWE-corrected  $p_{\text{cluster level}} < 0.05$ ).

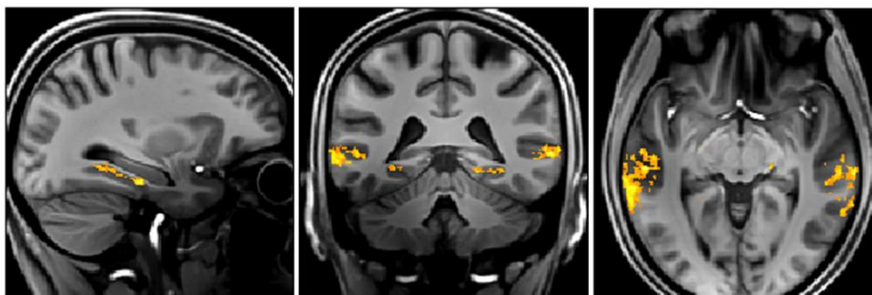
*Novelty:* The opposite contrast of higher activity for new than for old stimuli (PN > PO) was identified bilaterally in the parahippocampal and retrosplenial cortices of young adults (see Figure 26, and Table 6). Again, no brain regions were activated by this contrast for older adults. This was confirmed by a two-sample t-test showing bilateral PhC with higher activity in young compared to older adults (left:  $N_{\text{voxel}} = 249$ ,  $T_{\text{peak}} =$

4.964,  $p_{\text{peak}} < 0.001$ ,  $p_{\text{clusterFWE}} = 0.024$ ,  $p_{\text{clusterFDR}} = 0.007$ ; right:  $N_{\text{voxel}} = 356$ ,  $T_{\text{peak}} = 5.262$ ,  $p_{\text{peak}} < 0.001$ ,  $p_{\text{clusterFWE}} = 0.002$ ,  $p_{\text{clusterFDR}} = 0.001$ ).



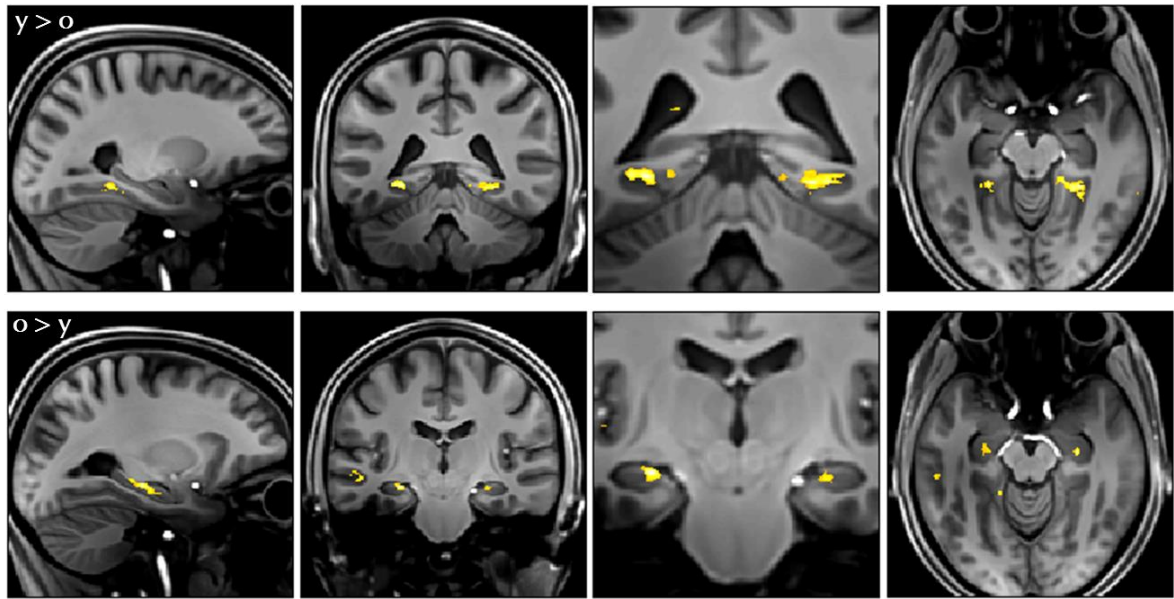
**Figure 26. Ch. 5 – Novelty effects in young adults only.** Greater activity for new than old stimuli (PN > PO) was observed in bilateral parahippocampal (left) and retrosplenial (right) cortices in young adults only (thresholded at  $p_{\text{voxel level}} < 0.001$ ;  $k > 100$  voxels; FWE-corrected  $p_{\text{cluster level}} < 0.05$ ).

*Visibility:* Brain regions that could differentiate between different levels of stimulus completeness included the superior temporal sulcus (STS) and parahippocampal cortex (PhC) bilaterally in both young and older adults (see Figure 27). Amygdala and lateral geniculate nucleus were also active in both age groups albeit different lateralization (see Table 6 and Table 7). Additionally, the hippocampus (including all subfields again) was involved in older adults only. However, this difference could not be confirmed by a two-sample t-test, as there were no significant clusters that were activated more by either age group.



**Figure 27. Ch. 5 – Visibility effects in both age groups.** Greater activity for full than partial stimuli (FO > PO) was observed in bilateral clusters in the parahippocampal cortices (PhC) as well as in the superior temporal sulcus (STS; thresholded at  $p_{\text{voxel level}} < 0.001$ ;  $k > 100$  voxels; FWE-corrected  $p_{\text{cluster level}} < 0.05$ ).

*Overall age effects:* On top of the age differences within specific contrasts, a repeated measures ANOVA (using the flexible factorial design) revealed a main effect of age over the whole task. Crucially, higher activity in young compared to older adults was observed in bilateral PhC, whereas higher activity in older compared to young adults was associated with bilateral hippocampus and left STS (see Table 8 and Figure 28). Both contrasts also survived FWE-correction on the whole-scan level (all voxels in the scanned slab) in bilateral PhC and bilateral hippocampus ( $p_{\text{whole-scan}} < 0.001$ ). Most interestingly, the hippocampal cluster mainly covered CA3 (see bottom panel in Figure 28).



**Figure 28. Ch. 5 – Age effects across the whole task.** (Top) Greater activity for young compared to older adults ( $y > o$ ) was observed in bilateral clusters in the parahippocampal cortices (PhC). (Bottom) Greater activity for older than young adults ( $o > y$ ) was observed in bilateral hippocampus (mainly CA3) and left superior temporal sulcus (STS); thresholded at  $p_{\text{voxel level}} < 0.001$ ;  $k > 100$  voxels; FWE-corrected  $p_{\text{cluster level}} < 0.05$ ). Hippocampus and PhC were also significant on the whole-scan level FWE-corrected ( $p_{\text{whole-scan}} < 0.001$ ).

**Table 6. Ch. 5 – Brain regions showing significant effects in young adults.**

contrast	region	L/R	$N_{\text{voxel}}$	peak $T$	cluster $p_{\text{FWE}}$
pattern completion (PNc > POc)	superior temporal sulcus	L	4199	6.67, 4.9, 4.44	0.000
erroneous pattern completion - new stimuli (PNc > PNf)	superior temporal sulcus	L	3999	6.08, 5.7, 4.73	0.000
		R	1772	4.9, 4.6, 4.31	0.000
retrieval (PO > PN)	hippocampus $\diamond$	L	2168	6.91, 6.31, 5.25	0.000
		R	966	7.23, 6.02, 5.45	0.000
	amygdala	L	777	6.08, 5.29	0.000
		R	293	5.26	0.003
novelty (PN > PO)	parahippocampal cortex $\diamond$	R	808	5.7, 4.95, 3.74	0.000
		L	383	5.42, 4.9	0.000
	retrosplenial cortex	L	236	6.09	0.014
		R	192	5.59	0.046
visibility (FO > PO)	superior temporal sulcus	L	13134	7.24, 7.12, 6.99	0.000
		R	7884	7.85, 7.25, 7.24	0.000
		R	197	5.20	0.038
		R	305	5.78	0.002
	parahippocampal cortex	R	1518	6.8, 6.69, 5.68	0.000
		R	457	5.94, 3.79	0.000
	L	350	5.24	0.001	
	amygdala	L	191	5.21, 5.14	0.045
	lateral geniculate nucleus	L	358	6.59	0.001
		L	276	7.44	0.004

Brain regions showing group-level activation (thresholded at  $p_{\text{voxel level}} < 0.001$ ;  $k > 100$  voxels;  $n = 20$ ). Only regions surviving cluster-level family-wise error (FWE) corrections are presented ( $p_{\text{cluster level}} < 0.05$ ). Several  $T$  values indicate multiple peak voxels within that cluster.  $\diamond$  indicates the region in that contrast was only significant in young adults. Hemisphere: L = left, R = right; conditions: FO = full old, PO = partial old, PN = partial new, c = correct, f = false. Coordinates are not presented as they are not meaningful beyond this sample, because a study-specific group template was used.

**Table 7. Ch. 5 – Brain regions showing significant effects in older adults.**

contrast	region	L/R	N <sub>voxel</sub>	peak <i>T</i>	cluster $p_{FWE}$	
pattern completion (PNc > POc)	superior temporal sulcus	L	1207	5.01, 4.63, 4.4	0.000	
		L	399	5.12, 3.4	0.000	
		L	262	5.36	0.005	
		R	192	4.58	0.033	
		R	178	4.91	0.050	
erroneous pattern completion - new stimuli (PNc > PNf)	superior temporal sulcus	L	5219	6.12, 5.72, 5.68	0.000	
		L	189	5.78	0.036	
		R	1269	4.76, 4.72	0.000	
		R	433	5.25, 4.17	0.000	
		R	327	5.1, 3.58	0.001	
erroneous pattern completion - old stimuli (PNc > POfo)	superior temporal sulcus $\diamond$	L	7530	5.94, 5.79, 5.77	0.000	
		L	418	5.0, 4.0	0.000	
		L	229	5.92	0.012	
		R	2297	5.02, 4.58, 4.3	0.000	
		R	1134	5.54, 5.09, 4.42	0.000	
		R	284	5.7, 3.77	0.003	
		R	351	5.52	0.001	
retrieval (PO > PN)	-					
novelty (PN > PO)	-					
visibility (FO > PO)	superior temporal sulcus	L	2911	7.45, 7.09, 6.67	0.000	
		L	287	5.24, 4.87	0.001	
		L	256	6.28, 4.84	0.002	
		R	476	5.81	0.000	
		R	344	6.88, 4.54	0.000	
		R	241	7.01	0.004	
		R	213	5.30	0.009	
		R	206	5.21	0.011	
		R	173	6.41, 3.93	0.032	
		hippocampus	R	1586	9.08, 6.5, 5.97	0.000
			L	550	6.66, 4.9	0.000
		parahippocampal cortex	R	1552	5.94, 5.49, 4.58	0.000
			L	1185	5.96, 5.43	0.000
		amygdala	R	574	7.07, 3.96	0.000
		lateral geniculate nucleus	L	242	6.04	0.004
	R	215	5.46	0.008		

Brain regions showing group-level activation (thresholded at  $p_{\text{voxel level}} < 0.001$ ;  $k > 100$  voxels;  $n = 19$ ). Only regions surviving cluster-level family-wise error (FWE) corrections are presented ( $p_{\text{cluster level}} < 0.05$ ). Several *T* values indicate multiple peak voxels within that cluster.  $\diamond$  indicates that region in that contrast was only significant in older adults. Hemisphere: L = left, R = right; conditions: FO = full old, PO = partial old, PN = partial new, c = correct, f = false, fo = false identified as old. Coordinates are not presented, as they are not meaningful beyond this sample, because a study-specific group template was used.

**Table 8. Ch. 5 – Brain regions showing significant effects between age groups across the whole task.**

contrast	region	L/R	N <sub>voxel</sub>	peak <i>T</i>	cluster $p_{FWE}$
young > old	parahippocampal cortex*	L	760	5.89	0.007
		R	1833	5.80, 4.35, 4.22	0.011
old > young	hippocampus (CA3)*	L	901	5.75, 4.61, 3.76	0.000
		R	415	5.55, 3.31	0.003
	superior temporal sulcus	L	1634	4.94, 4.85, 3.62	0.000
		L	359	4.90	0.007

Brain regions showing group-level activation (thresholded at  $p_{\text{voxel level}} < 0.001$ ;  $k > 100$  voxels;  $n = 39$ ). Only regions surviving cluster-level family-wise error (FWE) corrections are presented ( $p_{\text{cluster level}} < 0.05$ ). Several *T* values indicate multiple peak voxels within that cluster. \*indicates that the contrast was also significant when FWE-corrected on the whole-scan level (albeit in smaller clusters). Hemisphere: L = left, R = right. Coordinates are not presented as they are not meaningful beyond this sample, because a study-specific group template was used.

## 5.4. DISCUSSION

In summary, the behavioural data showed that recognition memory performance decreased with reduced stimulus completeness, older adults performed worse than young adults, and they had a higher tendency to pick familiar responses, which again replicated the previous findings (see chapters 2.3 and 3.3.1). Furthermore, both age groups made more errors identifying new stimuli as compared to learned stimuli. Here, participants evidently learned over the course of the experiment, but older adults learned less. Unfortunately, concomitant neuroimaging data assessed by multi-voxel pattern similarity were inconclusive, because the correlation results did not carry the information to distinguish between conditions. Univariate results nevertheless revealed a variety of regions involved in different aspects of the task, with regions in the memory network playing a prominent role. The superior temporal sulcus (STS) possibly stretching into lateral occipital complex (LOC) was consistently associated with contrasts likely engaging pattern completion processes. The STS together with the parahippocampal cortex (PhC) were also linked to visibility determined by stimulus completeness. Crucially, some prominent age effects could be identified: PhC was associated with novelty, and hippocampus with more general retrieval in young adults only. Most strikingly, independent of task condition older adults showed reduced PhC activity, and hippocampal hyperactivity mainly in CA3 in line with rodent findings (Wilson et al. 2005), and strongly supporting models of cognitive aging (Wilson et al. 2006).

First of all, parts of the previous behavioural findings could be replicated and provide further evidence that older adults resort to familiar responses more often than choosing new ones in line with computational models (Wilson et al. 2006). However, one finding was different from the previous studies in that young adults were also less accurate for new stimuli as compared to learned ones. Simultaneously, the variance in the young group's performance was also unusually high for the new stimuli (as compared to other conditions and previous versions of the task). Thus, here, young adults also had an increased tendency to choose familiar responses when presented with new items, minding that they were nevertheless more accurate than older adults. Arguably, this happens because in this version of the task participants never saw the



complete new stimuli making it harder if not impossible for them to form a coherent representation. Nevertheless, what this may show us instead is that participants were still learning during the test phase. This was in fact the case, as is now supported by the learning effects observed across sessions. More specifically, the general performance advantage of younger adults reported in the previous studies (see chapters 2.3 and 3.3.1) may reflect their superior and continued learning but which was mitigated here by the diminished information available for new stimuli. For example, in previous versions of the task it may have sufficed to see a new stimulus once in full to establish a sufficient representation and consequently identify its partial versions which was not possible in the current paradigm. Moreover, older adults may have been additionally impaired in encoding the new stimuli during the test phase in combination with their retrieval deficits. This interpretation would favour an added pattern separation account contributing to the observed age differences which is in line with the literature (Yassa et al. 2011b). While older adults were biased towards pattern completion during retrieval, they were also impaired in pattern separation during encoding (as predicted by Wilson et al. 2006). Although, the MIC has been designed to preferentially engage retrieval, simultaneous encoding was, of course, still possible. This is not a counter argument against the task design, it only highlights the necessity to investigate these processes in concert as they likely interact. Thus, although the MST favours encoding (respectively pattern separation) and the MIC favours retrieval (respectively pattern completion), both processes are probably acting together in each task (for discussion, see also Hunsaker and Kesner 2013). In addition, the scanner environment may also have contributed to the behavioural discrepancy, because the magnetic field can cause stronger dizziness and vertigo in 7T MRI which may result in a feeling of stress, nausea and fear impacting performance (Muehlhan et al. 2011). This may have affected the younger age-group more than the older adults in line with research on increased stress-hormone release in adolescents undergoing MR-scanning (Eatough et al. 2009).

With regard to the neuroimaging data, the absence of any multivariate pattern similarity effects may have several reasons. One possibility is that stimulus-specific representations could not be detected by the BOLD signal, although other studies have successfully used this method (e.g. Stokes et al. 2015; Kyle et al. 2015a). Alternatively, the patterns may not be stimulus-specific or the hippocampus may not code for these representations, but this has also been shown before (previous references, and e.g. Schlichting et al. 2014; Aly and Turk-Browne 2015). However, it should be noted that most studies show very low correlation values in the hippocampus ( $r < 0.05$ , e.g. in Copara et al. 2014; Kyle et al. 2015b), and some studies have also failed to identify hippocampus involvement in memory paradigms using pattern similarity. For example, one study employed an incidental encoding paradigm (continuous change detection of an overlaid fixation cross) with stimuli grouped into categories (faces, body-parts, objects, scenes) rather than manipulating single object similarities, and a surprise recognition test after scanning (LaRocque et al. 2013). Here, multi-voxel pattern similarity could not identify the hippocampus as a relevant structure in pattern separation, because Pearson correlations in the hippocampus did not differ between

stimuli or stimulus categories, but rather in the PrC and PhC. Thus, the hippocampus may not necessarily code for all information involved in memory. Furthermore, it is also possible that the current methods and specifically functional analysis with 7T MRI may not be advanced enough yet to detect these kinds of specific subcortical signal changes; e.g. preprocessing of these images is already considerably more difficult since standard algorithms of common software packages (SPM or FSL) often fail with 7T data in terms of coregistration or normalization. While these issues have been solved in this study (by using whole-brain EPIs, a combination of softwares, and more advanced tools like ANTs; see Methods 5.2.5), it is possible that processing of the functional data with GLMs is susceptible to similar difficulties albeit their detection may go unnoticed as there is no observable ground-truth comparable to checking the match of aligned images. A recent study also suggests that higher resolution in 7T is not necessarily beneficial, because a standard resolution of 2 mm isotropic voxels yielded better decoding results in visual cortex in comparison with data of up to 0.8 mm resolution (Sengupta et al. 2017). Furthermore, behavioural variability may have added extra noise to neural signals complicating the identification of performance- and stimulus-specific hippocampal involvement.

As for the univariate results, no "active" pattern completion processes could be identified. That is, no region showed more activity for correctly recognized learned partial images than for new stimuli. However, the superior temporal sulcus (STS) possibly stretching into lateral occipital complex (LOC) stood out in contrasts associated with pattern completion under a repetition suppression account. That is, when comparing activity for correctly recognized partial learned stimuli (completed) to correct identification of new partial stimuli (no pattern completion necessary), the STS showed less activity for the learned images consistent with adaptation (see Methods 5.2.5.3 for more detail on repetition suppression effects). These patterns nicely fit to a suggestion by Raymond Kesner and Edmund Rolls (Rolls and Kesner 2006; Kesner and Rolls 2015; Rolls 2013, 2016) that STS should be the region where recall is produced from the output of CA1 and ErC following hippocampal processing. Usually, the STS is reported to be involved in multisensory integration, motion, face and speech processing, but a review suggests that especially the posterior portion of STS (which was involved here) is not strictly functionally divided but rather supports different cognitive functions depending on task-dependent network connections involving MTL and frontal cortex (Hein and Knight 2008). STS is generally assumed to be part of the ventral stream in the hippocampal-cortical network involved in memory-guided behaviour (Ranganath and Ritchey 2012). Additionally, the identified cluster potentially stretched into LOC although this cannot be said with absolute certainty as this region is usually functionally defined (Malach et al. 1995). However, the findings are consistent with the literature on object recognition, where the LOC showed high activation to grid masks of objects (not unlike the masking used here) compared to scrambled masks where the order of visible pieces was distorted (Lerner et al. 2002). The LOC's role in object recognition has been strengthened in other studies (Grill-Spector et al. 2001; Rose et al. 2005) along with a particular susceptibility to adaptation effects (Kim et al.

2009). Furthermore, a recent study investigating pattern completion in multi-element engrams has also observed LOC involvement specifically for objects (Horner et al. 2015) pointing towards cortical reinstatement as a result of successful pattern completion. Although the stimuli used here were not objects but scenes, partial versions may have resembled several objects rather than a coherent scene. STS was also involved when pattern completion led to an erroneous result, i.e. when a stimulus was falsely recognized as a learned one. But only in older adults was STS associated with errors for learned images, which may be related to the particularly selective increase in these error rates for older adults (cf. Table 5). Overall, STS (and LOC) seemed to reflect the reinstatement of a learned stimulus independent of whether it was actually presented and more expressive of the concurrent behavioural outcome, which may be the result of successful pattern completion.

Interestingly, regions involved in more general retrieval and novelty could only be identified in young adults. Although these contrasts were defined independent of behaviour, decreased accuracy of older adults for all involved conditions may be related to the absence of observable effects. Yet for young adults, hippocampus and amygdala were involved in retrieval, i.e. higher activity for learned as opposed to new stimuli. Although unspecific, that is, independent of actual behavioural outcome, this may reflect a hippocampal retrieval process potentially linking cortical areas (Staresina et al. 2013b). At any rate, the general involvement of the hippocampus in memory retrieval has been demonstrated extensively (O'Reilly et al. 1998; Hasselmo and McClelland 1999; Squire 2004; Yonelinas et al. 2010) and the amygdala has also been attributed to memory processes in terms of reward discrimination (Gilbert and Kesner 2002), retrieval of conditioned taste aversion (Osorio-Gómez et al. 2017), and influencing the hippocampus in emotionally modulated retrieval (Leal et al. 2014; Zheng et al. 2017).

Next, parahippocampal and retrosplenial cortices were involved in novelty processing. Although the hippocampus has often been implicated in novelty working as a match/mismatch detector (Kumaran and Maguire 2007, 2009), other regions upstream including the PhC have also been found to contribute to novelty (Kafkas and Montaldi 2014), with one study even reporting a double dissociation between hippocampus and PhC supporting PhC's selective involvement in novelty (Howard et al. 2011). Retrosplenial cortex has been shown to be important for the detection of novel spatial arrangements, in line with its overall prominent role in spatial processing (for review, see Vann et al. 2009).

Further on, STS was also associated with visibility, that is, when a stimulus was shown in full it was more active than when the stimulus was only partially visible. Given that full stimuli were the easiest to identify, retrieval was most successful in these trials possibly explaining STS' stronger involvement in line with its supposed role in retrieval (Rolls 2013). The PhC was also associated with visibility consistent with its involvement in visuospatial processing. More specifically, PhC is involved in scene processing and has been found to produce more activation for full scenes as opposed to objects or close-up scenes (Henderson et al. 2008), which translates well to the current

observations on full (scene) and partial images (close-up like). Interestingly, visibility was also observed in the older adults' hippocampus (though not statistically different from young adults), potentially consistent with its role in scene perception with higher complexity (Zeidman and Maguire 2016).

Critically, the lack of observable effects for retrieval- and novelty-related activity, respectively in hippocampus and PhC in older adults is in line with numerous studies identifying the MTL as especially vulnerable to neurodegeneration with age (Jagust 2013). Moreover, age-related deficits in novelty detection have been frequently reported, e.g. in the discrimination of new environments in aged rats (Burke et al. 2010), or older humans (Kirasic 1991), or reflected in the absence of novelty-related event related potentials (Friedman 2000; Fjell and Walhovd 2005). Similarly, older adults often present with numerous retrieval deficits (see Introduction 1.4.2 for more examples), e.g. they are impaired in recalling where or in which context they experienced a certain event (Koen and Yonelinas 2016), or need more hints to trigger recall (Lindenberger and Mayr 2014). Most strikingly, I have identified behaviour-independent age differences across the whole task manifesting in reduced PhC-activity and CA3-hyperactivity in older adults. First of all, CA3-activity levels have been found to be elevated in aged rats across old and new environments (Wilson et al. 2005), which has since been suggested to play a primary role in altered hippocampal circuitry and function in aging (Wilson et al. 2006). There are hints from neuroimaging studies pointing into the same direction, where elevated DG/CA3 activity in older adults has been shown in comparisons of novel and lure items (using the MST; Yassa et al. 2011a, 2011b). However, these studies did not differentiate between DG and CA3 which likely show distinct activity patterns, and they were based on a very specific contrast assuming repetition suppression. Critically, in these studies, the age differences were directly tested within the DG/CA3 complex by extracting beta estimates and comparing their mean between age-groups. Here, CA3 was identified independent of conditions, and it stood out in testing age-effects across all scanned voxels. Thus, I provide strong evidence for overall hyperactivity in the CA3 of older adults when performing a recognition memory task, supporting models of cognitive aging that suggest an age-related bias towards pattern completion (Wilson et al. 2006). Secondly, the PhC showed reduced activity in older adults. Even though the hippocampus usually stands out in the report of age-related changes (Raz et al. 2005; Reagh et al. 2015; O'Shea et al. 2016), there are some studies reporting similar findings. For example, a study investigating scene memory found reduced PhC-activity in older adults alongside compensatory activity in inferior frontal regions when remembered and forgotten stimuli were compared (Gutchess et al. 2005). This could have been the case here too, however, the functional scanning sequence did not cover frontal regions. Another study found that volume in the posterior PhC was specifically affected in a group showing memory decline over the course of 12 years in contrast to participants with no memory decline (Burgmans et al. 2011). Even more, one study suggested that PhC volume may be a better biomarker than the hippocampus in discriminating between healthy aging, MCI and early Alzheimer's disease (AD) especially in its early stages (Echávarri et al.

2011). While I tried to exclude participants with starting cognitive impairment (using the MoCA), and the brains were checked for severe structural differences, I cannot eliminate the possibility that some of the participants may show early signs of AD as other imaging modalities sensitive to these brain changes were not included in the scanning protocol.

To summarize, the STS was identified as a place of reinstatement for retrieved partial images. Furthermore, the hippocampus played a role in differentiating learned from new images possibly reflecting an active retrieval mechanism, and the PhC was associated with novelty. Crucially, both novelty and retrieval contrasts were not observed in older adults suggesting that hippocampal and parahippocampal processing is impaired in aging. Unfortunately, specific contributions of individual hippocampal subfields could not be identified due to a variety of reasons discussed above. This may not be taken to suggest that the hippocampus is not involved in pattern separation and completion, but that it could not be specified given the diverse methodological constraints. It is of importance to note, that while the hippocampus was not observed in retrieval contrasts for older adults, the STS was associated with erroneous reinstatement for learned stimuli which was also reflected in higher error rates for these trials. Most importantly, this study provides strong evidence for CA3-hyperactivity and reduced PhC-activity in older adults.

## 5.5. CONTRIBUTIONS

Thomas Wolbers helped conceptualize the study and discussed analysis steps with me. The study was a cooperation with the DZNE site in Bonn, where all the testing was carried out. Jennifer Faber supported me in all administrative issues and communication, and she recruited and tested about a third of the participants. Tony Stöcker, Rüdiger Stirnberg and Daniel Brenner provided support with all MR-related issues and imaging protocols. Myung-Ho In (from the Leibniz Institute for Neurobiology Magdeburg) provided and installed the sequence for point spread function mapping. Anke Rühling was the scanning assistant.

Carla Bilsing recruited and tested most of the participants, analysed parts of the behavioural data and wrote up the results for her Bachelor's thesis in Psychology at the Otto-von-Guericke-University Magdeburg. Katharina Mamsch segmented the hippocampi of 75% of the young sample after careful training by me.

## 6. SEGMENTATION PROTOCOL

### Segmentation protocol for MTL subregions in 7T-MRI

As a result of my occupation with hippocampal subfields, and the need to cleanly differentiate them for accurate determination of their function, it became clear that current segmentation protocols are inconsistent with the most recent neuroanatomical findings. Even though I already applied a more valid border distinguishing CA3 and DG in my 7T study (see chapter 5.2.3), more discrepancies regarding the hippocampal head emerged later imposing the necessity to adjust the current sets of rules. Additionally, other MTL regions upstream of the hippocampus were also involved in the task and will deserve a closer look in the future.

The protocol as presented in the following chapter has been published in revised form according to reviewer suggestions:

Berron D\*, Vieweg P\*, Hochkeppeler A, Pluta JB, Ding S-L, Maass A, Luther A, Xie L, Das SR, Wolk DA, Wolbers T, Yushkevich PA\*, Düzel E\*, Wisse LEM\* (2017). A protocol for manual segmentation of medial temporal lobe subregions in 7 tesla MRI. *NeuroImage: Clinical* 15: 466-482. [*\*Denotes equal first and senior author contributions.*]

#### 6.1. INTRODUCTION

The human hippocampus and the adjacent medial temporal lobe (MTL) regions have been implicated in a number of cognitive functions including episodic memory (Eichenbaum et al., 2007), spatial navigation (Ekstrom et al. 2003; Epstein 2008; Wolbers and Büchel 2005) and perception (Aly et al. 2013; Graham et al. 2010; Lee et al. 2005a). At the same time, MTL regions are also affected by a number of pathological conditions such as depression (Huang et al. 2013), posttraumatic stress disorder (Wang et al. 2010), epilepsy (Bernasconi et al. 2003), and neurodegenerative diseases like Alzheimer's Disease (Dickerson et al. 2004; Du et al. 2007; Rusinek et al. 2004).

In vivo MRI research on the functional anatomy of the human hippocampus and the MTL has made considerable progress over the past years and novel neuroanatomical findings have increased our knowledge on subdivisions of the MTL (for a recent atlas, see Ding et al., 2016). More specifically, novel data became available on the boundaries of the subdivisions in the perirhinal cortex (PrC) - area 35 and area 36 - depending on sulcal patterns that differ between hemispheres in continuity and depth (Ding and Van Hoesen 2010). Likewise, the boundaries of subiculum (Sub) and CA1 in the hippocampal head (HH), have been shown to feature anatomical variations between individuals that depend on the number of hippocampal digitations (Ding and Van Hoesen 2015).

Until recently, a major barrier to further advances in research on the MTL was the sparse anatomical reference from histological studies to guide in vivo segmentation of

these regions. Often only a few slices from a small number of cases are presented in histological reference materials. As a result, there is limited information available about the location of subregion boundaries for in vivo segmentation protocols, especially with regard to anatomical subvariants. Additionally, most of the extant histological reference material is based on samples sectioned at orientations different from in vivo T2-weighted MR images, used for MTL subfield segmentation, that are typically obtained perpendicular to the long axis of the hippocampus (Yushkevich et al. 2015a). It is unclear how much this difference in orientation affects the relative location of the boundaries within the MTL and how well it translates to in vivo MR images, especially in the more complex head region of the hippocampus (these issues were also mentioned in (Wisse et al. 2016a). The aforementioned new histological study by Ding and Van Hoesen addresses this point as they presented data from 15 samples sectioned perpendicular to the long axis of the hippocampus, thereby matching commonly used MR images, and providing more than a single case to account for anatomical variations between brains (Ding and Van Hoesen 2015).

However, the delineation of these small MTL structures on in vivo MRI is also limited by the information available in commonly used MR images. T2-weighted 3T images with high in-plane resolution are generally used to delineate hippocampal subfields because of the visualization of the stratum radiatum lacunosum moleculare (SRLM), which appears as a thin dark band on these scans and can be used to define borders between some of the subfields. Such high in-plane resolution can often only be obtained at the cost of either lower signal-to-noise ratio or larger slice thickness, given the limited scan time available especially in clinical populations. This limits the precision of the measurements in-plane and along the long axis of the hippocampus. Recent developments at 3T, and the increased availability of ultra-high resolution MRI at 7T now allow for increased signal-to-noise ratio and resolution and thereby a more consistent visualization of internal features from slice to slice while maintaining a smaller slice thickness (of up to 1 mm) in a reasonable scan time.

Indeed, several segmentation protocols have been published for 7T (Boutet et al. 2014; Goubran et al. 2014; Maass et al. 2015, 2014; Parekh et al. 2015; Suthana et al. 2015; Wisse et al. 2012) leveraging the improved visualization for the distinction of small subfields such as the dentate gyrus (DG) and CA3 in the hippocampus (Parekh et al. 2015), the SRLM (Kerchner et al. 2012), and allowing for specific analyses of subregions of the ErC (Maass et al. 2015) and even entorhinal layers (Maass et al. 2014). However, most 7T protocols limit the segmentation to the hippocampal body, with the exception of Wisse et al. (2012) and Suthana et al. (2015). Additionally, most published 7T protocols have not reported inter- or intra-rater reliability. Aside from that, there is also a considerable number of hippocampal subfield segmentation protocols available at lower field strengths (Daugherty et al. 2015; La Joie et al. 2013; Malykhin et al. 2010; Mueller et al. 2007; Winterburn et al. 2013; Zeineh et al. 2001). However, none of these protocols – neither at 3 nor at 7T – have incorporated the anatomical variations dependent on hippocampal indentations (Ding and Van Hoesen 2015).

Manual segmentation protocols that include extrahippocampal regions PrC and PhC mainly exist for T1 weighted images with standard resolution at 3T MRI ((Feczko et al. 2009; Insausti et al. 1998; Kivisaari et al. 2013; Pruessner et al. 2002) but see (Suthana et al. 2015) and (Yushkevich et al. 2015b) for T2 weighted high-resolution MRI). The manual segmentation of the ErC and PrC is complex, as the anatomical boundaries change based on the depth of the collateral sulcus, which is further complicated by the variability of sulcal patterns across hemispheres. Only few protocols at 3T provide depth-specific rules (Insausti et al. 1998; Kivisaari et al. 2013) and even less consider different subvariants of sulcal patterns (Feczko et al. 2009). Crucially, no protocol has yet incorporated the more comprehensive and novel information on MTL anatomy (Ding et al. 2009; Ding and Van Hoesen 2010) while leveraging the high resolution at 7T to provide subvariant-specific and depth-dependent rules that account for the variability across and within subjects.

There is a large multi-investigator effort currently underway to develop a harmonized protocol for hippocampal subfields and extrahippocampal subregions (Wisse et al. 2016b), following the harmonized protocol for the total hippocampus (Apostolova et al. 2015; Frisoni et al. 2015). This harmonization effort was launched to overcome significant differences reported between extant segmentation protocols (Yushkevich et al. 2015a). However, the harmonization effort is currently aimed at 3T MRI (first limited to the hippocampal body, to be followed by expansion to the head and tail), and the protocol for 7T and extrahippocampal regions is not anticipated for several more years. The aim of this study was therefore to establish a segmentation protocol to manually delineate subregions in the parahippocampal gyrus as well as hippocampal subfields while leveraging the information available at 7T MRI images. We incorporated newly available information from histological studies and conferred with a neuroanatomist (S.-L.D.) to provide anatomically valid and reliable rules. Intra- and inter-rater reliability results are included for 22 hemispheres of younger adults. Additionally, this manuscript provides a comprehensive description of the segmentation of extrahippocampal regions, including detailed information on anatomical variation between subjects.

## 6.2. MATERIALS AND METHODS

### 6.2.1. PARTICIPANTS

Participants were included using baseline data from a study investigating the effects of physical exercise on the brain. Exclusion criteria were reports of regular sports activities that improve cardiovascular fitness as well as high physical activity levels. In addition, participants were screened for known metabolic disorders and neurological or psychiatric history, and excluded from further examination in case of incidents reported during history taking. Participants were recruited from the Otto-von-Guericke University campus in Magdeburg. Fifteen young and healthy individuals (16 hemispheres) were included from the baseline scan before any intervention. Seven additional subjects (8 hemispheres) were included after refining the rules for sulcus depth measurements (age range 19-32; mean age = 26, 12 female; see 1.1.1 and 6.3).



In total, we used 24 hemispheres of these subjects. All subjects gave informed and written consent for their participation in accordance with ethic and data security guidelines of the Otto-von-Guericke University Magdeburg. The study was approved by the local ethics committee.

---

### 6.2.2. WORKSHOP

In order to test the usability of the manual segmentation protocol, we hosted a segmentation workshop for 35 participants who were mostly novices (29 out of 35). The protocol was sent out four weeks prior to the workshop in combination with example MR images in order to give participants the opportunity to familiarize themselves with the segmentation approach. On site, we presented the protocol followed by an intensive hands-on session. From that, we used the given feedback and most commonly occurring problems to refine the protocol and improve comprehensibility for novice raters. This includes figures that provide a quick overview of the rules (Figure 37), more detailed annotations of the slice-by-slice plots (Figure 31, Figure 32, Figure 34, Figure 35) as well as supplemental material of cases with rare anatomical variants.

---

### 6.2.3. IMAGE ACQUISITION

Imaging data were collected at the Leibniz Institute for Neurobiology in Magdeburg on a 7T MR scanner (Siemens, Erlangen, Germany) with a 32-channel head coil (Nova Medical, Wilmington, MA). We acquired 22 partial turbo spin echo (TSE) T2-weighted images oriented orthogonal to the long axis of the hippocampus (in-plane resolution =  $0.44 \times 0.44$  mm, 55 slices, slice thickness = 1 mm, distance factor = 10%, TE = 76 ms, TR = 8000 ms, flip angle =  $60^\circ$ , FOV = 224mm, bandwidth = 155 Hz/Px, echo spacing = 15.1 ms, TSE factor = 9, echo trains per slice = 57). The slice thickness of 1 mm together with the 10% distance factor results in a distance of 1.1 mm between slices. Scan-time was 7:46 minutes.

---

### 6.2.4. SEGMENTATION SOFTWARE

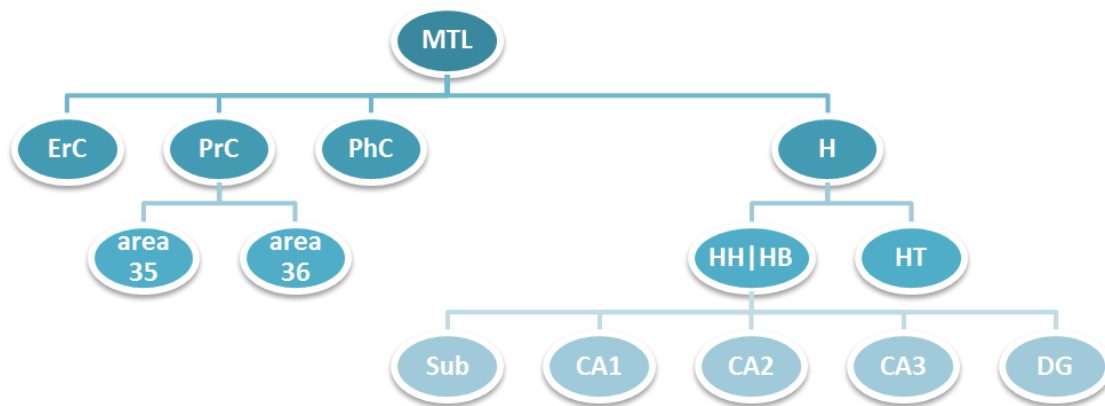
Structures were manually traced by two experienced raters (A.H. and A.L., see 1.1.1 for details) on oblique coronal slices using ITK-SNAP (Version 3.4; [www.itksnap.org](http://www.itksnap.org); Yushkevich et al., 2006). The images were adjusted for equivalent contrast range prior to segmentation (by capping the contrast curve at a maximum of 500). ITK-SNAP provides very useful features for implementing this protocol, i.e. an annotation tool for drawing lines and measuring distances.

---

### 6.2.5. MANUAL SEGMENTATION PROTOCOL

The protocol describes rules for manual segmentation of structures in the MTL in coronal MR images. The segmentation guidelines for the parahippocampal cortex (PhC), perirhinal cortex (PrC; area 35 and 36), entorhinal cortex (ErC) as well as the outer contours of the hippocampus are described in the first part (6.2.5.2), and further subdivision of the hippocampus into subfields are described in the second part (6.2.5.3;

for a segmentation hierarchy see Figure 29). Boundary rules are based on recent data from neuroanatomical atlases (Ding et al. 2016; Mai et al. 2015; Ding and Van Hoesen 2010, 2015). In this protocol we separately report neuroanatomical evidence and resulting rules which can be applied to MR images. Boundary rules are provided in millimeters in order to make the protocol applicable to scans of different resolution and facilitate comparisons with the neuroanatomical literature. The protocol is particularly focused on T2-weighted images acquired at 7T with  $0.44 \times 0.44 \text{ mm}^2$  in-plane resolution and 1 mm slice-thickness with 0.1 mm spacing. Some inner boundaries described in the section about hippocampal subfields, especially the boundaries of CA3 and DG that rely on the visualization of the endfolial pathway (Lim et al. 1997), are likely only applicable to 7T high-resolution T2 images. However, the described protocol could potentially also be applicable to other images that are acquired orthogonally to the long axis of the hippocampus with similar in-plane resolution and larger slice-thickness (e.g. 2 mm slice thickness).



**Figure 29: Segmentation hierarchy.** Segmentation of entorhinal cortex (ErC), area 35 and 36 of the perirhinal cortex (PrC), parahippocampal cortex (PhC) and the whole hippocampus separated into head (HH), body (HB) and tail (HT) are described in 6.2.5.2 (dark blue) and segmentation of hippocampal subfields is described in 6.2.5.3 (light blue).

### 6.2.5.1. ANATOMICAL LABELS USED IN THE PROTOCOL

In this protocol, we segment ErC, PrC, PhC and the hippocampus. We differentiate between area 35 and 36, which are frequently considered together as constituting the PrC in manual segmentation protocols (Duncan et al. 2014; Ekstrom et al. 2009; Olsen et al. 2013; Preston et al. 2010; Zeineh et al. 2001), except for Kivisaari et al. (2013) and Yushkevich et al. (Yushkevich et al. 2015b). However, these regions constitute different neuroanatomical parts of PrC (Ding and Van Hoesen 2010). Therefore, following the terminology of Ding and Van Hoesen, we refer to these regions as area 35 and 36. Note that these regions are slightly different from Brodmann areas 35 and 36 as the latter extend more posterior than area 35 and 36 in our study (for discussion, see Ding and Van Hoesen 2010). We note that area 35 roughly corresponds to the transentorhinal region (Braak and Braak 1991) and also to the medial PrC (Kivisaari et al. 2013). Detailed guidelines for hippocampal subfields involve the boundaries between the subiculum (Sub), CA fields 1-3 and the dentate gyrus (DG). Note that our segmentation of Sub includes subiculum proper, prosubiculum, presubiculum and parasubiculum (Ding 2013). Also, the DG here includes the hippocampal hilus or

region CA4, as these cannot be separated at this field strength. The SRLM is equally divided between its surrounding structures and not segmented separately. Hippocampal subfield segmentation encompasses the whole hippocampal head (HH) and body (HB) and is not performed in the tail (HT) because of the limited information with regard to the subfield boundaries in this region.

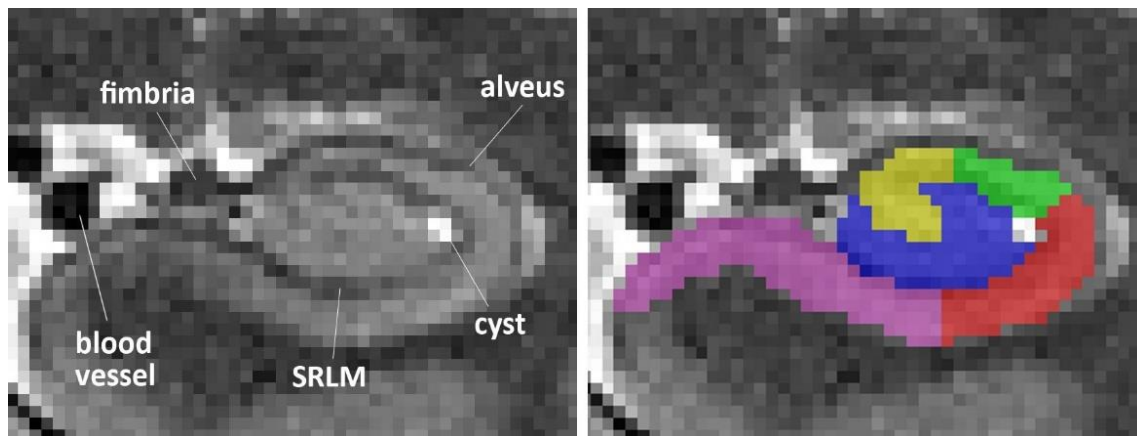
---

#### 6.2.5.2. HIPPOCAMPUS AND SUBREGIONS IN THE PARAHIPPOCAMPAL GYRUS

##### *Exclusions: alveus, fimbria, cerebrospinal fluid and blood vessels*

---

Fimbria and alveus as well as blood vessels, all appearing hypointense (see Figure 30) are excluded from anatomical masks (Wisse et al. 2012; Yushkevich et al. 2010). In general, the hippocampus is enclosed by white matter, visible as a hypointense line surrounding it. This line is spared from segmentation in this protocol. Additionally, there are several blood vessels within and close to the hippocampus. Both blood vessels and potential concomitant signal dropout should be excluded from the segmentation. Cerebrospinal fluid (CSF) and cysts appear hyperintense on T2-weighted MRI. Cysts, often located in the hippocampal sulcus (hippocampal fissure) at the ventrolateral flexion point of CA1 (van Veluw et al. 2013) are given a separate label. CSF - either surrounding the hippocampus or along a whole sulcus (e.g. hippocampal, uncal, collateral, occipito-temporal sulci) - are entirely excluded from the anatomical masks. CSF in sulci can be given a separate label as CSF (see explanation for transitions and labelling of the sulci in 6.2.5.3 below).



**Figure 30: Excluded structures in a coronal view.** Anterior hippocampal body slice from a T2 MRI scan including alveus, fimbria, SRLM, a blood vessel and a cyst in the ventrolateral flexion point of CA1 in the vestigial hippocampal sulcus.

#### *Hippocampal formation*

---

In the following we provide segmentation rules separately for the hippocampal head, body and tail. This is done to structure the following section rather than to construct independent masks of head, body and tail portions.

##### *Hippocampal head*

---

The anterior tip of the hippocampal head (HH) can be easily identified without additional landmarks (see Figure 31, HH0). Once the uncal sulcus can be followed from its fundus to the medial surface, the ErC becomes the inferior boundary of the HH,

which is segmented by connecting the most medial point of the white matter to the most medial point of the grey matter (see Figure 31, HH4; (Wisse et al. 2012). At the posterior end of the HH, the uncus separates from the hippocampus (see Figure 31, HH14). While it is still connected to the rest of the HH (via grey matter), the hippocampus is segmented as one structure (see Figure 31, HH13). Once the uncus is separated (e.g. only connected via the fimbria), the HH and uncus are segmented as separate structures in the coronal plane (see Figure 31, HH15).

### *Hippocampal body*

---

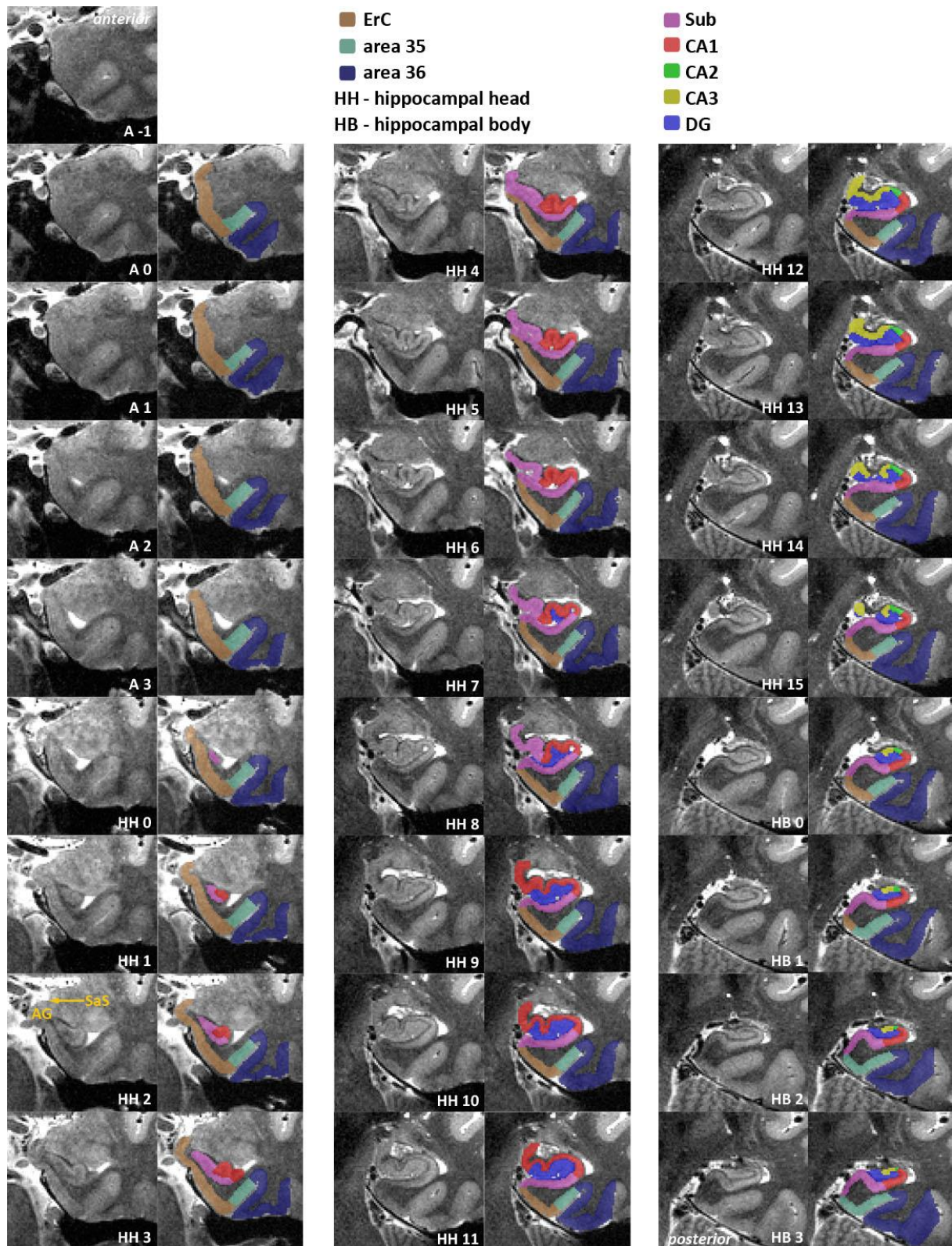
The hippocampal body (HB) begins when the uncus has disappeared (1 slice posterior to the uncus apex; see Figure 31, HB 0). White matter and CSF surround the HB superiorly, medially and laterally. The medial-inferior boundary of the HB is the connection of the most medial point of the white matter to the most medial part of the grey matter, where it successively borders ErC, area 35 and PhC (see Figure 31 HB 0 - HB 3, e.g. (Ding and Van Hoesen 2010). Sometimes, in more posterior slices, a small sulcus (the anterior tip of calcarine sulcus; CaS) appears medially between HB and PhG. In this case, the lateral and medial banks of the CaS are spared from segmentation (see Supplementary Figure 1). However, often the CaS only appears in HT. The HB is segmented as long as the inferior and superior colliculi are visible (medial butterfly-shaped structures; Wisse et al. 2016c). Segmentation does not stop before the colliculi have disappeared entirely. This rule has to be applied for each hemisphere separately (see Supplementary Figure 2).

### *Hippocampal tail*

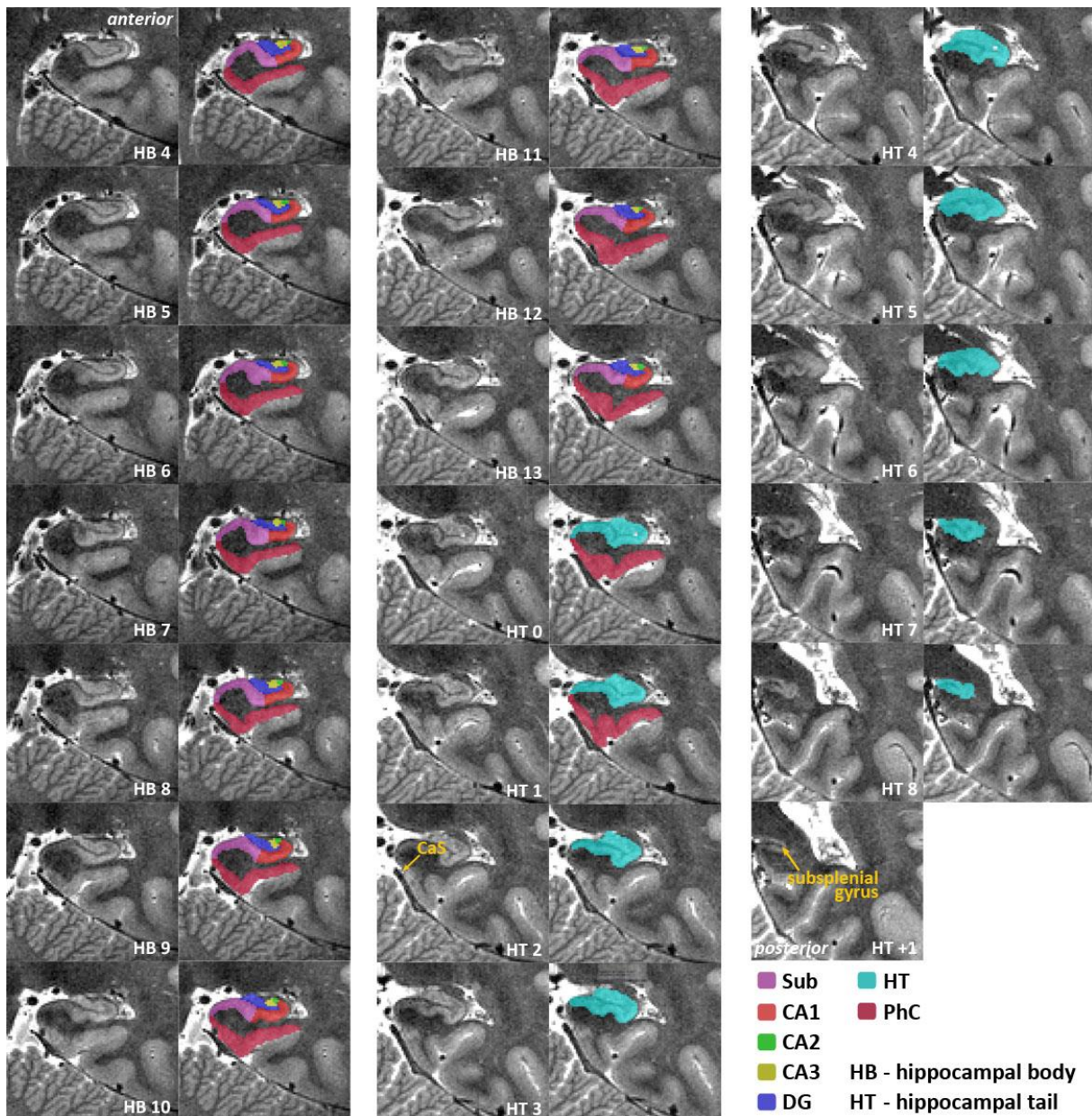
---

The hippocampal tail (HT) is a structure that is surrounded by white matter laterally, superiorly and ventrally. Most of these white matter structures are represented by alveus, fimbria and fornix, and are therefore excluded from segmentation. The medial-inferior boundary is constructed in the same way as that for the HB (see Figure 32, e.g. HT0-6). In more posterior slices, the HT (supero)laterally neighbors CSF in the trigone of the lateral ventricle (see Figure 32, e.g. HT3). The last slice of the hippocampus is the last slice where the HT is clearly visible (see Figure 32, HT11) which can also be checked on sagittal slices. It should be noted that at the very end of HT the hippocampus might medially blend with a gyrus, sometimes referred to as subsplenial gyrus. This gyrus is included in the hippocampal mask (Ding et al. 2016) until it is no longer connected to the rest of the hippocampal grey matter.

Note that different definitions of the body/tail border exist. Here, we chose the colliculi as they are easily identifiable, and are intended to provide a reliable posterior border for subfield segmentation.



**Figure 31: Slice-by-slice segmentation for a type 1 collateral sulcus (CS) – anterior part.** Slices are 1.1 mm apart. Included are entorhinal cortex (ErC; brown), perirhinal cortex (area 35 in mint green, area 36 in dark blue), subiculum (pink), CA1 (red), CA2 (green), CA3 (yellow) and dentate gyrus (blue). Shown in HH2, ErC covers the ambiens gyrus (AG) and superiorly ends at the semiannular sulcus (SaS). SaS constitutes the superior border of ErC and should be extrapolated to anterior slices when it cannot be identified there.



**Figure 32: Continuation of Figure 31 - Slice-by-slice segmentation for a type 1 collateral sulcus (CS) – posterior part.** Slices are 1.1 mm apart. Included are parahippocampal cortex (PhC; dark pink), subiculum (pink), CA1 (red), CA2 (green), CA3 (yellow), dentate gyrus (blue), and the hippocampal tail which, is not divided into subfields. In HT7, the subsplenial gyrus starts medially blending into the hippocampus. As soon as it is detached from the hippocampus, it is excluded from segmentation (HT+1). Delineation of PhC stops at the calcarine sulcus (CaS) in HT2.

### Entorhinal cortex

Segmentation of the ErC (as well as area 35 and area 36) begins 4.4 mm (= 4 slices here) anterior to the first slice of HH. That is, 4 slices have to be counted anterior to the hippocampus to define the starting slice. Although the ErC extends through most of the anterior temporal lobe (Ding and Van Hoesen 2010; Kivisaari et al. 2013) we chose this border because it is easily identifiable, and high-resolution structural imaging protocols often do not cover the entire anterior MTL. The superior border in anterior slices is the semiannular sulcus (Ding et al. 2016; Mai et al. 2015). Sometimes, this sulcus is not visible from the most anterior end of ErC, in which case it should be extrapolated from more posterior slices where it can be clearly identified (see Figure 31, HH2). The ErC

covers the ambient gyrus (AG; see Figure 31, HH0-3). Note that the ambient gyrus is made up of different subfields in an anterior-to-posterior direction. While the ambient gyrus is occupied by the ErC in more anterior slices (Insausti and Amaral 2012), it consists of Sub and CA1 in more posterior sections (Ding and Van Hoesen 2015). Moving posteriorly, at the point where the uncal sulcus can be followed from its fundus to the medial surface, Sub becomes the new superior border (see Figure 31, HH4). It is constructed by drawing a line from the most medial part of the white matter to the most medial part of the grey matter (Mueller et al. 2007; Wisse et al. 2012; Yushkevich et al. 2015b). This rule applies until the posterior end of ErC. The lateral border of ErC mainly consists of white matter. With respect to the inferomedial border, in some subjects CSF can be discerned between the ErC and the laterally located meninges (Xie et al. 2017, 2016). Therefore, bright voxels medial to the ErC have to be spared from the segmentation (see Figure 31, HH1-7). It should be noted that the intensity can depend on how much space there is between the meninges and the cortex. Sometimes these voxels appear slightly darker than CSF at other locations because of partial voluming with surrounding voxels. Inferolaterally, the ErC is bordered by area 35. This boundary is constructed at  $\frac{1}{4}$  of the longest expansion of CS (from edge to top of the grey matter) as the shortest connection between CS and white matter (see Figure 33). The only exception from this rule occurs when CS is less than 4 mm deep (very shallow CS); in that case the boundary between ErC and area 35 moves more lateral to the extension of the fundus of the CS. ErC disappears approximately 2 mm after the HH (Insausti and Amaral 2012). Segmentation of the ErC stops therefore after 2.2 mm (= 2 slices here) into the HB, i.e. after 2.2 mm posterior to the uncal apex (see Figure 31, HB0). The last slice of ErC serves as an intermediate step between ErC and the increasing size of area 35. Therefore, the lateral border of the ErC shifts by dividing ErC in half (Ding and Van Hoesen 2010; Insausti et al. 1998; Mai et al. 2015).

### *Perirhinal cortex*

---

Segmentation of area 35 and area 36 of the PrC is dependent on the sulcal pattern within the MTL – especially the collateral sulcal patterns are highly variable between brains but can also differ between hemispheres of the same brain. There are two main types of MTL anatomy – one deep CS (Type 1; 45 %), and a discontinuous CS, which can be divided into an anterior (CSa) and a posterior section (CSp) (Type II; 52%) (Ding and Van Hoesen 2010). CSp is usually longer and deeper than CSa. Studies have found a negative correlation of the depth of the CS and the depth of the occipito-temporal sulcus (OTS). In subjects with a shallow CS, the OTS is often deep and vice versa (see Figure 33; (Ding and Van Hoesen 2010). In some cases, the CS is bifurcated, i.e. it appears to have two conjoined sulci; the more medial sulcus is used here in this case (i.e. for evaluating the depth of CS). When it is difficult to identify the sulcal pattern in one slice, it is recommended to check in adjoining slices and interpolate to the difficult slices.

Given the differences in anatomy, different segmentation guidelines have to be applied for the different sulcal patterns as well as the depths of the CS. It is highly

recommended to define the sulcal pattern for each hemisphere before starting the manual tracing. The following descriptions are visualized in Figure 33.

### *Area 35*

Segmentation of area 35 starts at the same artificially chosen slice as ErC, i.e. 4.4 mm (= 4 slices) anterior to the first HH slice. Neuroanatomical atlases indicate that the posterior border of area 35 falls within 5 mm of the anterior portion of the HB. Segmentations therefore end 4.4 mm (= 4 slices) into HB, which is also 2.2 mm posterior to ErC (Ding and Van Hoesen 2010; Insausti et al. 1998). In the most posterior 2.2 mm, area 35 borders the Sub medially (see Figure 31, HB1-HB3); in all anterior slices it borders ErC. The superolateral and inferomedial borders are in accordance with those of ErC (e.g. white matter and CSF or meninges). The lateral border of area 35 depends on the depths of the sulci, and is measured from edge to fundus of the respective sulcus on each individual slice. For that purpose, the edges adjacent to the sulcus are connected via a tangent line. The depth of the sulcus is now measured from the middle of this line to the fundus of the sulcus (see Supplementary Figure 3A; also schematics in Figure 33). If the sulcus bends, the depth is measured in separate legs along the middle of the sulcus (see Supplementary Figure 3B).

#### *Very deep CS (>10 mm)*

Area 35 occupies the two middle fourths of the medial bank of the CS. Its lateral boundary with area 36 is constructed at  $\frac{3}{4}$  of the medial bank of the CS (see Figure 33).

#### *Deep CS (7-10 mm)*

From the border to ErC, area 35 occupies the remaining  $\frac{3}{4}$  of the medial bank of the CS (see Figure 33); i.e. from  $\frac{1}{4}$  of the medial bank up to the top of grey matter.

#### *Shallow CS (4-7 mm)*

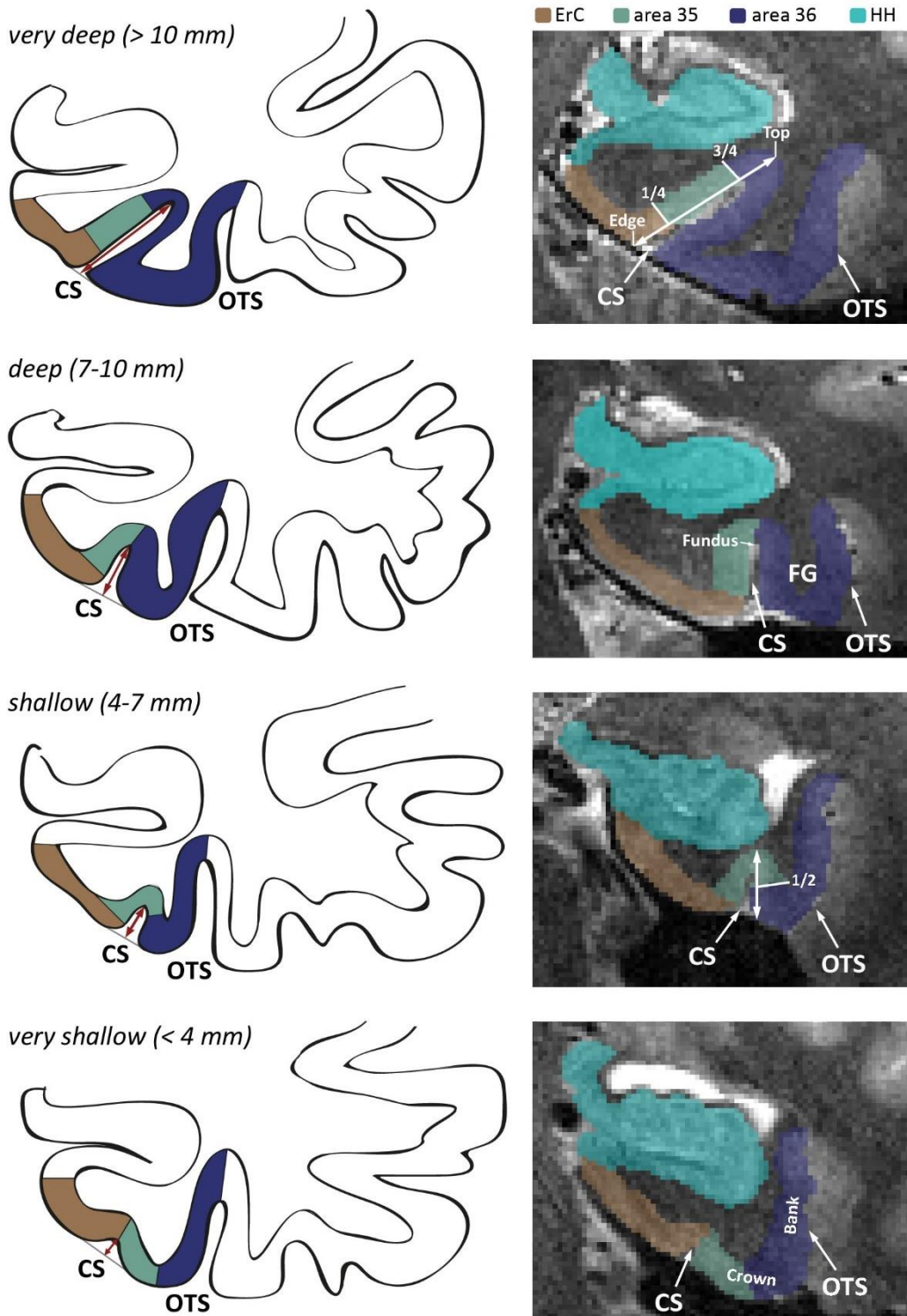
From the border to ErC, area 35 extends up to half of the lateral bank of the CS (see Figure 33).

#### *Very shallow CS (<4 mm)*

From the border to ErC, area 35 extends up to half of the crown of the fusiform gyrus (FG; see Figure 33).

When both CSa and CSp are visible on the same slice, the lateral boundary of area 35 is constructed at half of the crown between the two CS (see Figure 34, HH4). As soon as the CSa has disappeared, the same depth rules apply to CSp as shown in Figure 33. A decision tree can be used to facilitate the necessary decisions (see Supplementary Figure 4). The relationship of area 35 to area 36 length in histological studies roughly resembles a 1:3 ratio. Our rules are designed in order to approximate this ratio.





**Figure 33: Different depths of the collateral sulcus (CS) with respective segmentation rules applied.** Sulcus depth is measured from edge to fundus of CS as indicated by the red arrows. Edge, fundus, crown and bank are indicated for easy anatomical descriptions of the gyral and sulcal patterns. Quartiles for segmentation rules are defined by measuring the full extent of grey matter from edge to top along the respective bank as indicated by the white two-sided arrows in the images on the right. Entorhinal cortex (brown) ends laterally at  $\frac{1}{4}$  of the grey matter bank medial to CS, when CS is deeper than 4 mm. For very shallow CS (< 4mm), entorhinal cortex covers the whole medial bank of CS and ends at the extension of the fundus of CS. Segmentation rules for area 35 (green) change depending on the depth of CS: very deep – area 35 covers the middle part from  $\frac{1}{4}$  to  $\frac{3}{4}$  of the grey matter bank medial to CS; deep – area 35 covers the whole superior  $\frac{3}{4}$  of the grey matter bank medial to CS; shallow - area 35 extends up to half of the lateral bank of CS; very shallow - area 35 extends up to half of the crown of the fusiform gyrus (FG). Area 36 (blue) directly neighbours area 35 laterally, and extends towards the entire bank medial to occipitotemporal sulcus (OTS). The hippocampal head (HH) is depicted in turquoise.

## *Area 36*

---

Segmentation of area 36 is done in the same slices as area 35, that is, starting 4.4 mm anterior to the first HH slice, and ending 2.2 mm posterior to ErC. Area 36 directly borders area 35, thus its medial boundary depends on the different sulcal patterns described for area 35. Its lateral border is defined by the next lateral sulcus – the OTS. This border extends previous protocols (Feczko et al. 2009; Insausti et al. 1998; Kivisaari et al. 2013; Pruessner et al. 2002) and is specifically based on Ding and van Hoesen (2010). It is constructed by following the longest expansion of OTS, from medial edge to top of the grey matter, thereby including the whole medial bank of OTS (see Figure 33). It should be noted here that OTS is very variable, i.e. it can be bifurcated, or there could be two OTS. In these cases, the more medial OTS should be used as the border (see Supplementary Figure 5). Generally, OTS is rather deep and shows a reciprocal relationship with CS (see section 6.2.5.3 on area 35 for more detail), and thus can be differentiated from other small sulci that sometimes appear in-between CS and OTS, e.g. the mid-fusiform sulcus in posterior slices (for reference, see Ding et al. 2016) in Supplementary Figure 6.

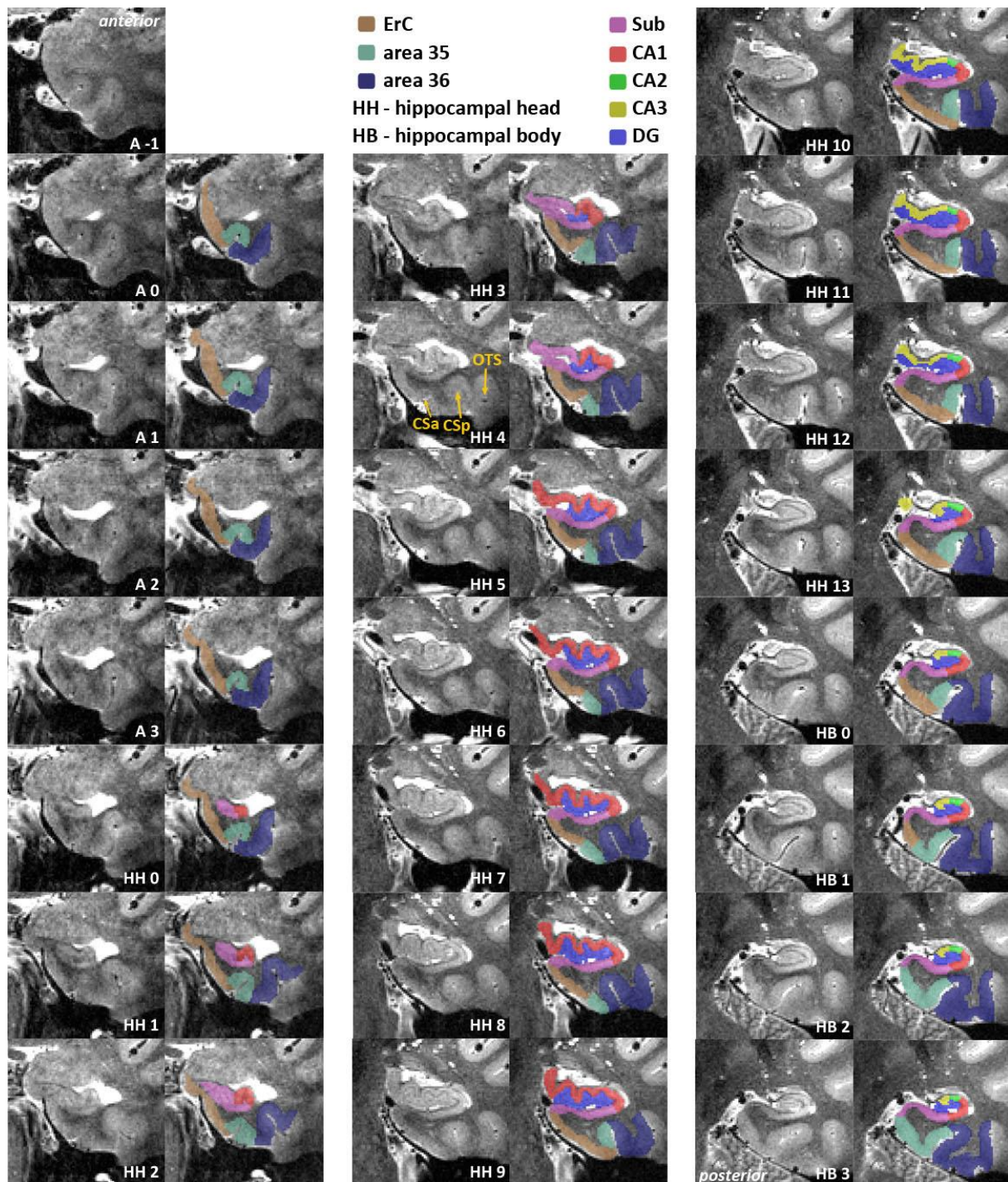
Sometimes, another small sulcus, the rhinal sulcus, is visible in very anterior slices. Generally, it is medial to CS and more shallow; it often 'travels' up the CS (see Supplementary Figure 7). In very rare cases, the rhinal sulcus can be separate from CS so far posterior that it affects segmentation. That is, when the rhinal sulcus is separate and visible on the medial cortical surface, the rules change in a way as if one were to substitute the CS with the rhinal sulcus and the OTS with the CS. The boundaries follow the same depth rules as above but are applied to the rhinal sulcus. The very lateral border of area 36 is now the CS and not the OTS. When the rhinal sulcus disappears, area 35 and area 36 change to the usual patterns.

## *Parahippocampal cortex*

---

Anteriorly, the segmentation of the PhC directly adjoins the posterior end of area 35. Thus, it begins 5.5 mm (= 5 slices) posterior to the uncus apex (see Figure 31, HB4; (Ding and Van Hoesen 2010; Insausti et al. 1998)). As with area 35, the PhC has a medial-superior boundary with Sub (see Figure 31, starting HB3). The superolateral and inferomedial borders are in accordance with those of ErC (e.g. white matter and CSF). The lateral boundary is the fundus of the CS extended to the top of grey matter. Posterior regions of the parahippocampal and fusiform gyrus include areas TH, TL and TF (Ding and Van Hoesen 2010). While TH and TL cover regions in the parahippocampal gyrus, TF occupies parts of the fusiform gyrus. The PhC in this study only covers temporal areas TH and TL, but not TF (Ding and Van Hoesen 2010; confer Ding et al. 2016). Segmentation stops when the anterior tip of the CaS appears medially - a small sulcus that mostly folds in a superior-to-inferior direction (see Figure 32, HT2; Supplementary Figure 1). Since little anatomical literature is available on the PhC, we based this decision on Song-Lin Ding's expertise annotating this region in histology samples, and because the CaS can be reliably identified in every subject. There is another small sulcus lateral to CaS, the newly discovered parahippocampal-ligul

sulcus (PhligS; (Ding et al. 2016); see Supplementary Figure 8), which would be a better indicator of PhC's borders. However, it was not possible to reliably distinguish this sulcus in every subject's MRI and we therefore chose to use the CaS as a landmark.



**Figure 34: Slice-by-slice segmentation for a type II collateral sulcus (CS) – anterior part.** Slices are 1.1 mm apart. Included are entorhinal cortex (ErC; brown), perirhinal cortex (area 35 in mint green, area 36 in dark blue), subiculum (pink), CA1 (red), CA2 (green), CA3 (yellow) and dentate gyrus (blue). HH4 is an example of a transition slice between anterior (CSa) and posterior CS (CSp) and the corresponding segmentation of area 35. The occipitotemporal sulcus (OTS) establishes the lateral border of area 36.

### *Transitions and labeling of the sulci*

---

In order to maintain smooth transitions between slices that resemble the anatomy more closely we introduce transitions. Whenever there are sudden changes from one rule to the other, or sudden appearances of anatomical structures, one intermediate slice serves as a transition. Thus, the last slice where the anatomy fulfills the criteria of one rule serves as a transition slice to the next rule by applying an intermediate step in the middle between both rules (e.g. see Figure 31, HB1; Figure 34, HH0 and HB1). This procedure should be used in the following cases: (1) when ErC ends posteriorly, (2) when only CSa changes to only CSp without both being visible on the same slice (see also Supplementary Figure 9) and (3) when the OTS “jumps” (appears/disappears) from one slice to the next. An optional additional label for the CS and OTS can be added to facilitate thickness measurements using automated tools (Yushkevich et al. 2015b). In case the CSF within the sulci is visible, these voxels can be labelled as sulcus. If the sulcus is not completely visible, there are usually some hints to it, such as an indentation on the inferior portion or a patch of CSF in the middle. If the sulcus cannot be identified, it can be estimated based on the thickness of the medial and lateral grey matter banks on surrounding slices. Inferring the sulcus in this way ensures that all voxels labelled as sulcus have adjacent edges (that is: not 1 voxel thick diagonal). In addition, if the gyri around CS and OTS touch, i.e. if no white matter in-between their grey matter banks is visible, the line of voxels in the middle between the two sulci should be artificially excluded from segmentation to allow meaningful thickness measurements. Alternatively, if the separation of the two banks can be inferred from the surrounding slices, a voxel line approximating that separation should be used instead.

---

#### **6.2.5.3. HIPPOCAMPAL SUBFIELDS**

These guidelines are mostly based on ex-vivo parcellations by Ding & van Hoesen (Ding and Van Hoesen 2015), and on comparative, additional information derived from other publications, such as the Mai atlas (Mai et al. 2015) and the protocol from Wisse et al. (2012).

#### *Sub and CA1 segmentation starts*

---

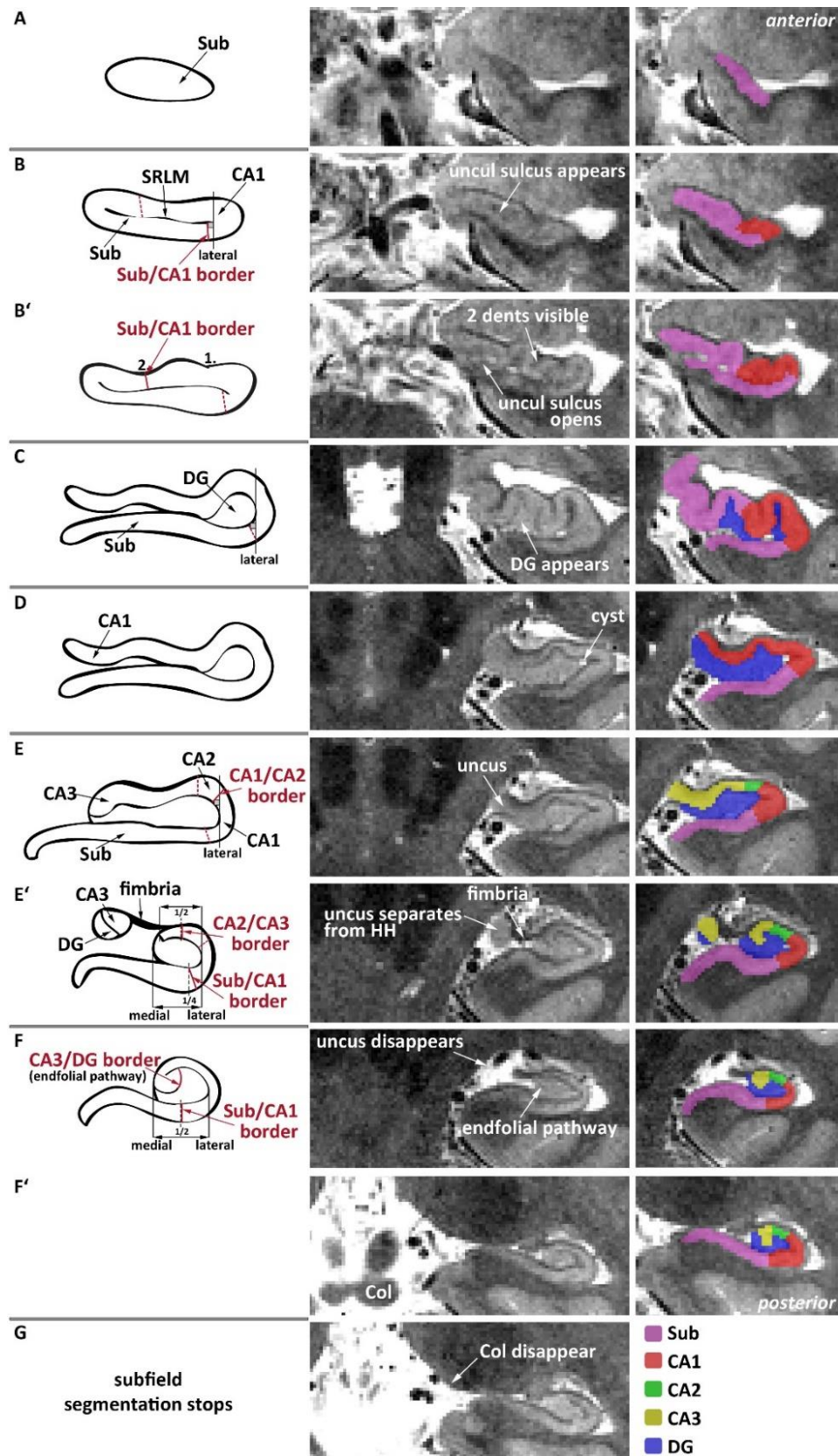
Mostly, the first anterior slice of the HH appears as one structure. Sub is then assigned to all of it (Figure 35A). Approximately 1-2 mm posterior to that, a hypointense line appears (i.e. uncal sulcus/SRLM; (Ding and Van Hoesen 2015)) dividing the hippocampus into a superior and an inferior part and shaping the hippocampus similar to a lip (Figure 35B). From here, the SRLM is equally divided between the regions it separates unless it is only 1 voxel wide, in which case it is segmented such that it always belongs to the superiorly located structure. Also at this point, the segmentation of CA1 starts. The guidelines can be more readily understood by looking at Figure 35B and B'. The inferior boundary (i.e. on the “lower lip”) between Sub and CA1 is an orthogonal line to the longitudinal Sub. It is positioned by finding the most lateral voxel of the SRLM, moving to the next medial one, and is constructed there from inner to outer side of the structure. The superior boundary between Sub and CA1 is extrapolated

from a posterior slice where the digitation of the HH can be clearly identified, i.e. when the “upper lip” has at least two dents (Figure 35B’). At the second indentation counted from lateral to medial, a straight line is constructed orthogonal to the structure, and copied to anterior slices. Posteriorly, the border is positioned at this same indentation on each individual slice. This border closely approximates what is observed in the hippocampal subvariants with two and three indentations, as described by Ding et al. (Ding and Van Hoesen 2015). Once the uncus sulcus opens, the separation of CA1 and Sub continues along the uncus sulcus (MR image in Figure 35B’). This may coincide with the appearance of DG, although it may also occur slightly more posterior.

### *DG segmentation starts*

---

When DG appears and does not yet stretch to the most lateral extension of the uncus sulcus (confer Figure 34, HH4), the inferior boundary between Sub and CA1 is constructed exactly like before. If, however, DG extends towards the most lateral point of the uncus sulcus (confer Figure 34, HH5), the reference point changes from one voxel medial from the most lateral SRLM to the most lateral DG voxel (Figure 35C). It is crucial to not confuse DG with cysts (which are brighter). However, if there is a cyst within DG that establishes the most lateral border, the cyst is used instead of DG to identify the CA1/Sub border (Figure 35D). Based on Ding and Van Hoesen (Ding and Van Hoesen 2015), the superior part of the subiculum disappears 1.2-1.8 mm after the appearance of the DG. We therefore chose to end segmentation of the superior part of the subiculum 2.2 mm (= 2 slices) after the first appearance of DG and this portion is then occupied by CA1 (Figure 35D). These borders are identified in the same way again on all following slices, although they often just stay the same as on the previous slices. From here, the SRLM is equally divided if thicker than 1 voxel, and otherwise segmented so it always belongs to the outer structure (i.e. CA1/Sub/etc., but not DG). It should be noted that contrary to the white matter surrounding the hippocampus, the hypointense line on the superior side of Sub is always included in the segmentation as it consists of the molecular layer of the Sub. Additionally, the inferior side of Sub is prone to signal drop-out due to the crossing perforant path; therefore voxels of intermediate intensity on the inferior side of Sub should be included because a very conservative visual segmentation of only the brightest voxels might result in an underestimation of Sub (see Bronen and Cheung 1991; Wisse et al. 2016c).



**Figure 35: Rules for hippocampal subfield segmentation shown on the relevant slices from anterior to posterior.** Schematic descriptions of all rules are depicted in the first column. Specific rule changes or new borders are indicated in red. Dashed lines are used, when the rule in question is inferred from another slice; e.g. the inferior Sub/CA1 border is defined in B, but the superior Sub/CA1 border is defined in B' and extrapolated anteriorly. The relevant anatomical changes are indicated by white labels and arrows in the middle column, e.g. when the uncus separates from the hippocampal body (HB), or the colliculi (Col) disappear. The resulting segmentation is shown in the last column; subiculum (Sub) in pink, CA1 in red, CA2 in green, CA3 in yellow and dentate gyrus (DG) in blue.

### *CA2 and CA3 segmentation starts*

---

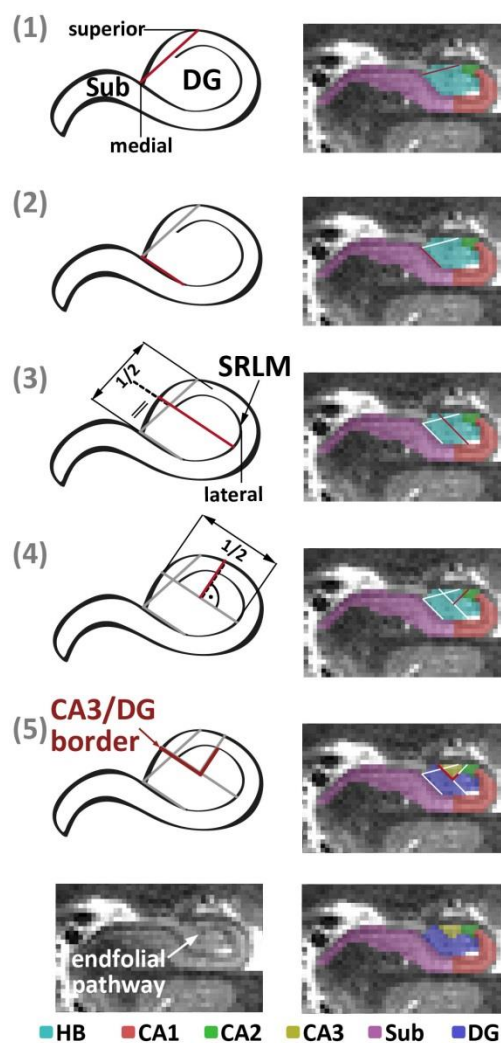
Neuroanatomical data indicate that the anterior border of CA3 in the head falls within 3-5.4 mm relative to the start of the head (Ding and Van Hoesen 2015). The segmentation of CA2 and CA3 therefore begins in the last 4.4 mm (= 4 slices) of the HH. Although CA2 generally appears before CA3 (Ding and Van Hoesen, 2015), there is limited information on the exact distance between the two and on potential differences between subjects. Therefore, we chose to start segmenting CA2 at the same slice as CA3. Additionally, CA2 and CA3 show an alternating pattern in the most anterior slices; we chose to simplify this and count all medial grey matter towards CA3. Although we realize that we may count some portions of CA2 towards CA1 or CA3, we chose for these simplifications to achieve high reliability. Again, it might help to consider Figure 35E and E' alongside this description. The border between CA1 and CA2 is constructed orthogonal to the CA structures at one voxel medial of the lateral boundary of DG; this is identical to the determination of the previous CA1/Sub border rule only on the superior instead of the inferior side. As in the previous section, if there is a cyst within DG that establishes the most lateral border, the cyst is used instead of DG to identify the CA1/CA2 border. The next step is to identify the point where the uncus separates from hippocampus. In some cases, only the fimbria is attached to both (Figure 35E'). The border between CA2 and CA3 is extrapolated from that slice to more anterior slices (to include the last 4.4 mm of HH). It is constructed halfway between the most medial point of the CA fields and the most lateral point of DG (it can therefore only be determined after the medial border of CA3 is determined). For all posterior slices, this border is determined slice by slice as a line orthogonal to the structure. The detached uncus is defined as CA3 unless there is a hypointense line, which can be used to differentiate between CA3 superiorly and DG inferiorly (Figure 35E')(Duvernoy et al. 2013). The Sub/CA1 border also changes within the last 4.4 mm of HH (Ding and Van Hoesen 2015). As soon as the uncus separates from hippocampus, the new border is marked at 1/4 from most lateral DG to most medial hippocampus proper. This line is extrapolated anteriorly to include the last 4 mm of HH (Figure 35E'). In the HB, i.e. when the uncus has disappeared, this border shifts to 1/2 from most lateral DG to most medial hippocampus proper (Figure 35F). This boundary is identified in the same way on all posterior slices.

### *CA3 and DG differentiation*

---

Unique to our protocol is the delineation of CA3 and DG. Depending on image quality and resolution, we propose two different rules. Both rules apply to the whole HB and the most posterior HH slices where the uncus is only connected via the fimbria. Many protocols have defined everything on the inner side of the SRLM as DG. Based on ex-vivo segmentations (Ding 2015; Mai et al. 2015) and the better contrast of T2 images and higher resolution of 7T imaging, the visualization of the endfolial pathway is possible, which can be used to more clearly differentiate between CA3 and DG (see the full HB segmentations in Figure 31; also in (Parekh et al. 2015; Wisse et al. 2016c). The endfolial pathway is followed from the medial edge of CA3 towards the point where it

intersects the SRLM. All voxels that lie supero-medially to this line belong to CA3 (see Figure 35F). However, if the endfolial pathway is not identifiable, or the aim of the research project is a comparison of groups where the endfolial pathway cannot be reliably distinguished in one group, an approximation can be achieved pursuing the following rules alongside Figure 36. First, construct a line from the middle and most superior part of the hippocampus to the medial edge of DG touching Sub (Figure 36-1). From the latter point, draw a line laterally along the dark band until Sub starts curving (Figure 36-2). Then, compose a line parallel to this which centrally intersects the first line (Figure 36-3) between the outermost extent of hippocampus until it intersects the SRLM. On the halfway point construct an orthogonal line towards the superior SRLM and close CA3 infero-laterally (Figure 36-4). All voxels lying superiorly to those lines belong to CA3 (Figure 36-5).



**Figure 36: Heuristic rules for separation of DG and CA3 if the endfolial pathway is not visible.** (1) construct a line from middle most superior part of the hippocampus to medial DG touching Sub; (2) from that point, draw a line laterally along the dark band until Sub starts curving, (3) parallel to this intersect line 1 centrally between SRLM and outermost extent of the hippocampus proper, ; (4) centrally intersect line 3 orthogonally; (5) CA3 assigned to voxels superior to lines 3 and 4. Applied rules are shown in the lower panel; unspecific hippocampal body (HB) in turquoise, subiculum (Sub) in pink, CA1 in red, CA2 in green, CA3 in yellow and dentate gyrus (DG) in blue. Compare the right panels for visual segmentation based on the endfolial pathway on the same slice.



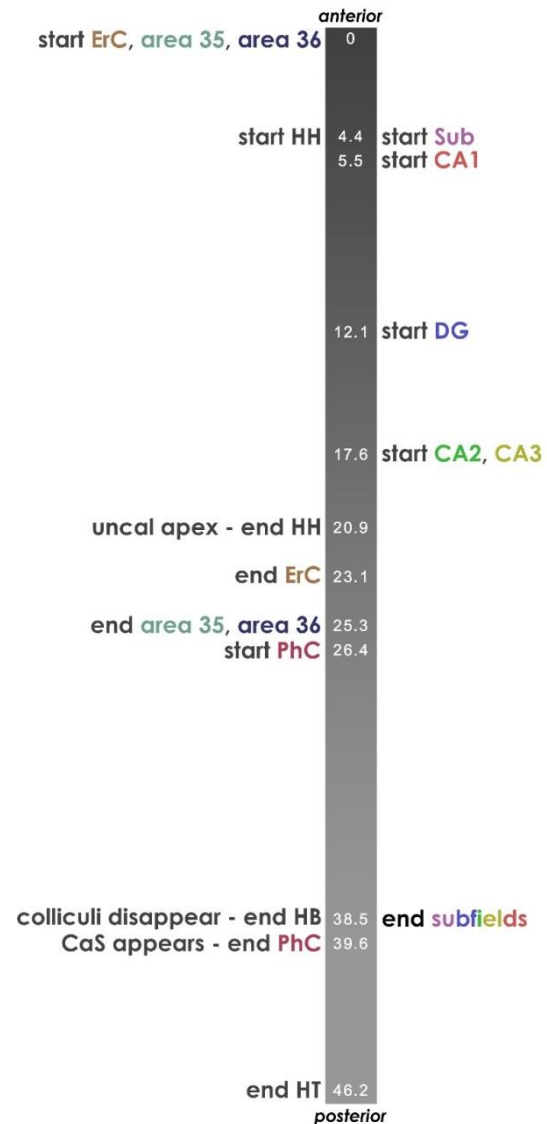
### Subfield segmentation ends

As described for the end of HB above, subfield segmentation stops when the colliculi (see Figure 35F') have disappeared entirely. This rule applies hemisphere-specific (Figure 35G; also Supplementary Figure 2). Afterwards, manual subfield segmentation is no longer reliable.

#### 6.2.5.4. GENERAL ADVICE FOR MANUAL SEGMENTATION

Segmentation of all regions is accomplished by tracing along white-to-grey matter boundaries, and several hypointense lines. These lines are not always continuous; we therefore recommend attempting smooth curvature even if the hypointense lines are discontinuous. Additionally, switching back and forth between coronal slices should ensure smooth transitions between slices, and avoid sudden jumps between regions. This is most important along the SRLM between Sub and CA1 in HH, at the endfolial pathway between CA3 and DG in HB, and for better identification of the sulcal pattern in PrC. Furthermore, special care is needed when measuring the depth of CS, because only slight variations can lead to different rule sets being required, i.e. at 7 mm rules for a shallow sulcus apply and at 7.1 mm rules for deep sulci apply (see Supplementary Figure 3, see rules for area 35 in 6.2.5.3; for impact of incorrect measurements see also results 6.3.1.2).

As shown above, there are many cross-references between areas, therefore we recommend defining certain key decision points prior to segmentation (see Figure 37). For example, the beginning and end of HH are needed as a reference for the start and end of ErC and areas 35 and 36. Additionally, we advise to check the full segmentation at the end (for a checklist, see Supplementary Figure 10).



**Figure 37: Exemplary segmentation profile.** This anterior-to-posterior axis (i.e. along the longitudinal axis of the hippocampus) illustrates the key decision points of this protocol (numbers indicate mm distance from the first anterior slice). Extrahippocampal regions and hippocampal head (HH), body (HB) and tail (HT) divisions are on the left; hippocampal subfields are depicted on the right. The start and end of each structure are depicted in the same color; often they depend on certain landmarks, e.g. the start and end of HH are used as a reference for the occurrence of the entorhinal cortex (ErC), area 35 and area 36, and the parahippocampal cortex (PhC). Most of these points are variable between brains but usually fall into a similar range as shown here. We recommend identifying these points prior to segmentation.

---

### 6.2.6. STATISTICAL ANALYSES

Two experienced raters (A.H. and A.L.) traced all subregions in the same 16 hemispheres independently (8 left and 8 right, 8 type I and 8 type II CS patterns). Both raters have each segmented around 40 subjects with the current rules prior to reliability testing. During that time, they met twice a week to discuss difficult cases, rule exceptions and to implement rule changes (e.g. the very shallow CS category was only introduced after specific feedback from S.-L.D.). From 14 subjects only one hemisphere was included, but from one subject two hemispheres were included to reach an equal number of type I and II CS patterns. In addition all subregions were segmented for a second time by one rater (A.H.) after 4 weeks.

The intra-rater reliability was assessed in 16 hemispheres in terms of relative overlap between the two time-points using the Dice similarity index (DSI; Dice 1945). The DSI was calculated for each MTL subregion. The consistency of volume measurements within one rater was assessed using intraclass correlation coefficients (ICC) using SPSS 22 (IBM SPSS Statistics for Macintosh, Version 22.0. Armonk, NY: IBM Corp.). The ICC variant that measured absolute agreement under a 2-way mixed ANOVA model was used (ICC(3); Shrout and Fleiss 1979).

The agreement of both raters was assessed in terms of relative overlap using the DSI and was calculated as before. The consistency of volume measurements between both raters was assessed using the ICC. This time, the ICC variant that measured absolute agreement under a 2-way random ANOVA model was used (ICC(2); Shrout and Fleiss 1979). Due to the low ICC values for area 35 in type II CS patterns in the first inter-rater reliability analysis, 8 additional hemispheres with type II CS were segmented by both raters, after sulcus depth measurements had been made more concrete in the protocol following careful evaluation of the mismatches encountered during the first round.

Average subregion volumes (mean and standard deviation) were calculated for the final 16 hemispheres for both raters (i.e. the 8 type I hemispheres and the 8 type II hemispheres from the second iteration after refinement of depth measurement).

---

## 6.3. RESULTS

---

### RELIABILITY

---

#### 6.3.1.1. INTRA-RATER RELIABILITY

Table 9 shows the intra-rater reliability of a single rater (A.H.) for 16 hemispheres (all from the first iteration). Almost all DSI values were above 0.9. Regions that were smaller and more complicated, such as CA2, CA3 and area 35, showed slightly lower values but were still over 0.85. ICCs were all over 0.95 with the exception of CA3 which was at 0.78 which may be explained by the more difficult but anatomically valid separation from DG along the endfolial pathway.

**Table 9. Intra-rater reliability for all subregions.**

	DSI (mean ± SD)	ICC
ErC	0.91 ± 0.01	0.98
area 35	0.88 ± 0.02	0.97
area 36	0.91 ± 0.02	0.96
PhC	0.93 ± 0.03	0.99
CA1	0.91 ± 0.02	0.98
CA2	0.87 ± 0.05	0.97
CA3	0.85 ± 0.03	0.78
DG	0.90 ± 0.02	0.98
Sub	0.92 ± 0.02	0.95
Hippocampus total	0.96 ± 0.01	0.97

Dice similarity coefficient (DSI) and intraclass-correlation coefficient (ICC) in 16 hemispheres from 15 subjects. CA = cornu ammonis, DG = dentate gyrus, Sub = subiculum, HT = hippocampal tail, ErC = entorhinal cortex, PhC = parahippocampal cortex

### 6.3.1.2. INTER-RATER RELIABILITY

In a first analysis, almost all DSI values were above 0.84. The DSI for smaller, more complicated regions CA2, CA3 and area 35 was slightly lower, though still over 0.77. Similarly, ICCs were above 0.87 for almost all subregions. The ICCs for DG and subiculum were slightly lower though still over 0.76. However, the ICC for area 35 was 0.68, and 0.47 for CA3. Since this number was discrepant from the remaining values, all segmentations were checked to find out whether the rules for CA3 were unclear and could therefore not be reliably implemented. An error was found in one subject by one rater. In this subject, the number of head slices in which CA3 was segmented was miscounted. As the last head slice was correctly identified and implemented for other labels depending on the most posterior head slice, we therefore concluded that this was a counting error rather than misinterpretation of the image or unclarity in the segmentation protocol with regard to CA3. This error was corrected and the ICC increased to 0.78.

In addition, we tested the inter-rater reliability of the heuristic rule for CA3 (see Figure 36). ICC was 0.78, and DSI was 0.79 indicating that the heuristic rule could be applied as reliably as using the endfolial pathway. On top of that, we compared the overlap between the two rules, therefore calculating an inter-rule DSI for all 32 hemispheres segmented by the two raters which was  $0.63 \pm 0.09$ . It should be noted here, that the upper limit for these DSI values are the DSI values of the inter-rater reliability for the two different sets of rules, i.e. with 100% rule overlap the DSI would be 0.79. Volume comparisons revealed that using the heuristic rule slightly underestimates CA3 (0.24 mL compared to 0.31 mL) and overestimates DG (1.0 mL compared to 0.93 mL) as compared to a separation at the endfolial pathway. This confirms that the heuristic rule is a good alternative to the anatomical landmark. However, if divergent from the endfolial pathway, in the majority of the slices CA3 will be underestimated, similar to all current segmentation protocols for CA3.

As an additional exploratory experiment, we performed reliability analyses of area 35 and 36 *separately* for the type I and type II sulcal variants. For area 36 the DSI was similar for the two types,  $0.87 \pm 0.02$  for type I and  $0.86 \pm 0.04$  for type II (see Table 10). The ICC for area 36 for type II was 0.99, higher than 0.84 for type I. For area 35, the DSI for type I was slightly higher than type II;  $0.84 \pm 0.06$  vs.  $0.78 \pm 0.07$ , but the difference was more notable for ICC of 0.87 for type I and -0.12 for type II. Although the ICC for area 35 is higher for type I as compared to type II, the absolute difference, or ‘measurement error’, between the two raters is similar for both sulcal pattern types (mean absolute difference: type I: 0.06 mL, type II: 0.07 mL) while the range of volumes for type II is only a third of the range of type I (range type I: 0.49-0.81 mL, type II: 0.59-0.72 mL). Thus, the absolute difference between the raters relative to the normal variation in the population (the range) for type II is larger than for type I. This is further illustrated in Bland-Altman plots in Supplementary Figure 11. Additionally, these plots show that neither rater had a bias as the differences lie around 0.

Further inspection of the segmentation of area 35 type II cases revealed that a difference in segmentation between the two raters mainly resulted from measuring the sulcal depth. A small difference in sulcal depth, as can be seen in Figure 33, can lead to a different segmentation rule. We therefore refined the segmentation protocol with regard to the sulcal depth measurements (see Supplementary Figure 3). Following this refinement of the protocol, 8 new type II hemispheres were segmented by both raters. The results of the reliability analyses for area 35 and 36 for these new type II hemispheres are presented in Table 10. ICC and DSI for ErC were 0.86 and 0.87, for area 35 they were 0.83 and 0.9, and for area 36 0.88 for both.

**Table 10. Inter-rater reliability in type 1 and 2 CS patterns separately.**

	DSI (mean $\pm$ SD)			ICC		
	Type I	Type II	Type II*	Type I	Type II	Type II*
<b>ErC</b>	$0.88 \pm 0.02$	$0.87 \pm 0.03$	$0.86 \pm 0.02$	0.94	0.80	0.87
<b>area 35</b>	$0.84 \pm 0.06$	$0.78 \pm 0.07$	$0.83 \pm 0.04$	0.87	-0.12	0.90
<b>area 36</b>	$0.87 \pm 0.02$	$0.86 \pm 0.05$	$0.88 \pm 0.03$	0.84	0.99	0.88

Dice similarity coefficient (DSI) and intraclass-correlation coefficient (ICC) for 8 hemispheres in each category. ErC = entorhinal cortex

\*Results from a second inter-rater reliability analysis following refinement of segmentation rules

As a result, the ICC and DSI for the combined type I and II cases improved to over 0.84 for all three regions. Table 11 shows the final results of all subregions for the comparison of both raters (i.e. average of the 8 type I hemispheres and the 8 type II hemispheres from the second iteration).

**Table 11. Inter-rater reliability for all subregions.**

	DSI (mean ± SD)	ICC
ErC	0.87 ± 0.02	0.94
area 35	0.84 ± 0.05	0.87
area 36	0.87 ± 0.02	0.88
PhC	0.86 ± 0.12	0.94
CA1	0.84 ± 0.04	0.89
CA2	0.81 ± 0.06	0.92
CA3	0.78 ± 0.05   *0.79 ± 0.05	0.76   *0.78
DG	0.86 ± 0.03	0.76
Sub	0.85 ± 0.04	0.78
Hippocampus total	0.94 ± 0.01	0.98

Dice similarity coefficient (DSI) and intraclass-correlation coefficient (ICC) in 16 hemispheres from 15 subjects. CA = cornu ammonis, DG = dentate gyrus, Sub = subiculum, HT = hippocampal tail, ErC = entorhinal cortex, PhC = parahippocampal cortex  
 \*Values after correction of CA3 segmentation on one subject for one rater.

### 6.3.2. VOLUMES IN COMPARISON TO ANATOMY

Mean volumes across both raters are shown in Table 12. Due to very different segmentation schemes for PrC as well as PhC, we did not compare volumes to earlier studies (Insausti et al. 1998; Pruessner et al. 2002). Volumes of hippocampal subfields are compared to earlier studies that used manual segmentation procedures at 7T (Wisse et al. 2016c), automated approaches at 3T (Iglesias et al. 2015; Yushkevich et al. 2015b) and histological techniques (Simić et al. 1997). The comparison of the mean volumes found in the current study with the other manually derived volumes from Wisse et al. (2016c) highlights the changes in the recent protocol. While CA1 has less volume compared to earlier estimations, subiculum shows an increase in volume. This probably relates to the new rule of subiculum segmentation in the hippocampal head as well as the new boundary between CA1 and subiculum in the hippocampal body. On the other hand, CA3 has increased while DG has reduced volume compared to Wisse et al. (2016c). Again, this most probably highlights our new rule which follows the endfolial pathway to separate DG and CA3 more accurately. Using this approach the portion of CA3 that folds into the DG is also segmented as CA3, which should result in an increase in CA3 volume. Additionally, hippocampal subfield volumes obtained with the current protocol approximate those obtained from post mortem studies (Iglesias et al. 2015; Simić et al. 1997), especially for the subfields in the hippocampus proper.

Table 12. Volumes of all regions in ml in comparison to other studies.

	Current study	Simic et al. (1997) <sup>1</sup>	Iglesias et al. (2015) <sup>2</sup>	Yushkevich et al. (2015) <sup>3</sup>	Wisse et al. (2016c) <sup>4</sup>
ErC	0.99 ± 0.2	-	-	-	0.53
area 35	0.64 ± 0.11	-	-	-	-
area 36	2.22 ± 0.39	-	-	-	-
PhC	0.58 ± 0.24	-	-	-	-
Sub	1.07 ± 0.16	∅	0.64*	0.34	0.63
CA1	0.82 ± 0.15	0.64	0.52	1.25	1.48
CA2	0.07 ± 0.02	-	-	0.018	0.07
CA3	0.17 ± 0.02	-	-	0.067	0.12
CA2&3	0.24	0.14	0.18	0.085	0.19
DG(&CA4)	0.50 ± 0.09	0.31	0.46	0.76	0.80
Hippocampus total	3.16 ± 0.40	1.54	2.26	2.44	3.1

CA = cornu ammonis, DG = dentate gyrus, Sub = subiculum, ErC = entorhinal cortex, PhC = parahippocampal cortex. We provide standard deviations for the data from the current study.

<sup>1</sup>Data derived from table 2, 'Normal'; <sup>2</sup>Data derived from table 3, 'Ex vivo atlas'; <sup>3</sup>Data derived from table 6, 'ASHS', mean of right and left side; <sup>4</sup>Data derived from table 1, 'Manual segmentation', mean of right and left side.

∅not shown because only entails subiculum and prosubiculum

\*values for parasubiculum, presubiculum and subiculum were summed up from table 3

## 6.4. DISCUSSION

We have developed and tested a new protocol for manual segmentation of the entorhinal cortex, perirhinal cortex (distinguishing area 35 and 36), parahippocampal cortex, and hippocampus as well as its subfields including subiculum, CA1, CA2, CA3, and dentate gyrus, in-vivo at 0.44x0.44 mm in-plane resolution using 7T MRI. We showed that our protocol had an intra-rater reliability ICC higher than 0.95, except CA3 (0.78) and DSI higher than 0.85 and an inter-rater reliability ICC higher than 0.76 and DSI higher than 0.81, except CA3 (>0.78) for all regions in young adults. The strengths of the protocol are outlined as follows.

First, we leveraged recent developments in neuroanatomy. This has enabled us to incorporate more distinct rules as previously known. Chiefly, subdivisions in HH and HB (Ding and Van Hoesen 2015), PrC (Ding and Van Hoesen 2010) and PhC (Ding et al. 2016) have substantially extended earlier work as they provide more details on the order of appearance and location of the subregions and additionally provide information on between-subject variability in some of the regions, which were incorporated in the current segmentation protocol, e.g. location of subregions in relation to different numbers of digitations in the hippocampal head and pattern of the collateral sulcus. Also note, that sectioning in one article (Ding and Van Hoesen 2015) was done perpendicular to the long axis of the hippocampus making it more comparable to the commonly used T2-weighted images. The other two atlases (Ding and Van Hoesen 2010; Ding et al. 2016) were based on histology data sectioned in a coronal plane. However, no histological data on extrahippocampal regions is currently

available with slices perpendicular to the long axis of the hippocampus. Additionally, earlier protocols mostly collapsed across subregions of the PrC instead of differentiating between medial and lateral parts (Insausti et al. 1998; Pruessner et al. 2002) but see (Kivisaari et al. 2013; Yushkevich et al. 2015b). We have extended that framework by differentiating more specifically between area 35 and 36 using available data from neuroanatomy (Ding and Van Hoesen 2010). Another example is the PhC, where studies have mostly included the posterior PhG up to CS across the whole length of the hippocampal tail because of lack of a well-established boundary (Pruessner et al. 2002; Yushkevich et al. 2015a). However, a recent histological atlas (Ding et al. 2016) disentangles the subdivisions of the posterior PhC. That is, the posterior PhC not only consists of areas TH and TL of the PhG, but also area TF of the fusiform gyrus. In addition, areas TL and TF extend further posteriorly than area TH, which in most cases disappears (replaced with ventral visual area V2) after the shallow parahippocampal-ligular sulcus (PhligS) appears. Although the newly identified PhligS would be anatomically the most valid landmark, it could not be identified reliably in all subjects. We observed that the anterior part of the CaS can serve as a landmark in close proximity to the PhligS, which can be distinguished reliably. The CaS is a rather conservative border and leads to an exclusion of a portion of posterior PhC. However, as this posterior portion also consists of visual area V2, its exclusion may benefit the study of parahippocampal function. This fine-tuning of the segmentation protocol with more detailed information on the borders and anatomical variability may further facilitate research on memory such as different memory pathways in the MTL (Das et al. 2015; Ranganath and Ritchey 2012; Reagh and Yassa 2014a). Additionally, if the protocol is validated in older populations, it may facilitate research on ageing and neurodegenerative diseases. For example, early stages of tau pathology in Alzheimer's Disease constitute especially in the transentorhinal region and the entorhinal cortex (Braak and Braak 1991; Ding et al. 2009). The transentorhinal region as described by Braak and Braak corresponds roughly to area 35 in the recent protocol. Therefore, a detailed volumetry of these regions that closely follows the anatomy becomes critical to detect early disease effects in volume and regional thickness measures (Wolk et al. 2017; Xie et al. 2017).

A second strength of this protocol is that we used ultra-high resolution 7T MRI, which enabled us to get more fine-grained images, and allowed for a more detailed delineation of smaller structures. In particular, the delineation of hippocampal subfields in the head as well as the visual distinction between CA3 and DG in the body benefit from the higher resolution. As can be seen from Figure 31 and Figure 34, the appearance of hippocampal head and presence of subfields can change drastically from slice to slice. The thinner slices obtained at 7T allow us to establish more precise segmentation rules for the hippocampal head – that is, determining the distance between the appearance of subfields in the order of 1 mm rather than the more frequently reported thickness of 2 mm. In addition, it likely also allows for a more reliable segmentation. This may add value when investigating diseases or cognitive functions for which the anterior portion of the hippocampus is proposed to be

specifically important (Poppenk et al. 2013; Sahay and Hen 2007). Additionally, we propose the use of the endfolial pathway, a white matter band aligned with the actual border of CA3, to separate CA3 from the DG in the hippocampal body in populations where it is visible. This accurate distinction of DG and CA3 may enable functional MRI studies to dissociate the contributions of DG and CA3, because they are assumed to be involved in different cognitive processes (Neunuebel and Knierim 2014). Although some of our rules are still geometrical in nature, the rules follow neuroanatomy more closely and take between-subject variability into account where possible. Additionally, most rules are independent of the in-plane orientation of the MTL; that is, most boundaries are drawn perpendicular to the structure rather than that they are based on the image orientation.

Thirdly, we aimed to develop a protocol that is easy to apply. Therefore, we included comprehensive slice-by-slice plots of high-resolution images that show the application of the rules along the full longitudinal axis for the most prevalent sulcal patterns – a continuous type I as well as a discontinuous type II CS (Ding and Van Hoesen 2010). In addition, we provide practical segmentation tips, a checklist for segmentation (Supplementary Figure 10) and schematic descriptions of the rules throughout the protocol as well as a decision tree for the segmentation of area 35 and examples of some difficult cases in the supplemental material. To further facilitate the understanding of our rules, we incorporated specific feedback from a workshop on our protocol in Magdeburg. One of our main aims was to understand the difficulties that novice raters encounter while learning to apply the protocol rules. During the workshop we identified the most common difficulties and adjusted the protocol accordingly. For example, it became clear that the frequent cross-referencing to certain anatomical structures (e.g. uncal apex) was difficult to follow. Therefore, we included recommendations in what order to approach segmentation (exemplified in Figure 37).

The intra-rater reliability showed that the protocol could be reliably applied across different time-points with DSI values higher than 0.85 and ICC values higher than 0.95 with the exception of 0.78 for CA3. The latter is probably due to the fact that we are using a more complex separation along the endfolial pathway, or using the heuristic rule. Analyses of the reliability between two raters showed that we were able to apply this protocol in a consistent manner, with almost all DSI values above 0.84 and almost all ICC values over 0.89. Even for smaller regions and for regions, such as area 35, for which the segmentation protocol is more difficult to accommodate anatomical variants, the ICC and DSI were reasonable (ICC over 0.68 for area 35 and over 0.76 for the other regions and DSI over 0.77), showing that these smaller and more complicated regions can be segmented with reasonable reliability. It should be noted that the ICC for the DG and CA3 was also slightly lower which can be explained by the more complex separation as already discussed for the intra-rater reliability. The high DSI values in general are encouraging for the application of this protocol to functional MRI studies as spatial overlap is most important in this context.

In relation to other studies, the ICC values and DSI values reported here are well within the range of previously reported reliability values (Bonnici et al. 2012a; de Flores et al.



2015a; Lee et al. 2014; Palombo et al. 2013; Prasad et al. 2004; Yushkevich et al. 2010; Goubran et al. 2014; Mueller et al. 2007; Winterburn et al. 2013). Although some other studies reported slightly higher values for CA1 (Shing et al. 2011; Lee et al. 2014; Yushkevich et al. 2010) or subiculum (Travis et al. 2014; de Flores et al. 2015a), our protocol includes more specific rules and may be more complicated. Additionally, the reliability for small regions such as CA2 and CA3 are among the highest reported in the literature. The ICCs for ErC, area 36 and PhC are also well in the range of previously reported reliability values (Feczko et al. 2009; Pruessner et al. 2002). The ICC value for area 35 was below the reliability estimates of earlier protocols (combining area 35 and 36). When splitting up the group in the two types of sulcal patterns, it became clear that this lower ICC value for area 35 was mainly driven by the type II variant. Importantly, our aim to incorporate the findings from histological studies (Ding and Van Hoesen 2010) and match anatomy as closely as possible resulted in a slightly more detailed protocol with segmentation rules dependent on sulcal depth measurements. Small differences in sulcal depth measurement could result in different segmentation rules especially in the type II variant. After initial evaluation of the results, we therefore refined the guidelines for sulcal depth measurements (see Supplementary Figure 3). A second reliability test in eight new type II hemispheres revealed an improved DSI of 0.83 and ICC of 0.90 which is similar compared to type I hemispheres. Although a learning effect could have affected the reliability measures, it seems unlikely that this played a large role as both raters had already segmented 40 subjects before the initial reliability test and this second reliability test was performed in 8 new hemispheres. These results indicate that with the refined segmentation protocol also a challenging region such as area 35 in the type II variant can be segmented reliably.

There are also limitations to the current study. First, by focusing specifically on anatomical validity and accounting for anatomical variability as much as possible, the resulting protocol is more elaborate and time-consuming compared to earlier approaches. However, we made considerable efforts to explain the protocol and make it understandable to novice raters. Additionally, we are planning to build an automatic segmentation of this protocol using ASHS (Yushkevich et al. 2015b). A second limitation is that although we tried to match anatomy as closely as possible, for some of the borders we still use heuristic rules to improve reliability of the protocol. This means that portions of subregions may be included in the labels of adjacent subregions. Third, our protocol was mainly based on work from Ding et al. and might not be in agreement with work from other neuroanatomists. However, our protocol is roughly consistent with other neuroanatomical references (Duvernoy et al. 2013; Insausti et al. 1998; Mai et al. 2015), and the volumes as obtained by the current protocol approximate the volumes from post mortem studies, as shown in Table 4 (Iglesias et al. 2015; Simić et al. 1997). The volume of the subiculum is slightly larger as compared to Iglesias et al., which may be due to the difference in age between the current study and the post mortem studies as subiculum volume is suggested to be affected by age (La Joie et al. 2010). It should be noted though that differences exist between these and other references in terms of nomenclature, for example the existence of the 'prosubiculum'

and perhaps also in the placement of certain boundaries. Fourth, although we embrace the possibilities provided by higher resolution, we are fully aware that not all researchers have access to 7T, which may limit the applicability of the current protocol. In order to facilitate application to 3T protocols, we provided all segmentation rules in millimeters and have included heuristic rules to guide segmentation of CA3 and DG when the anatomical landmark, the endfolial pathway, is not visible. Indeed we are currently trying out this segmentation protocol in a set of older adults and patients with MCI for whom a high resolution  $0.4 \times 0.4 \times 1.2 \text{ mm}^3$  T2-weighted MRI was obtained. Of note, although most studies on MTL subregions are using 3T imaging protocols, 7T might play a more prominent role in the future with an increasing number of sites with access to a 7T scanner. For example, recently the European Ultrahigh-Field Imaging Network for Neurodegenerative Diseases ([EUFIND](#)) was founded with the aim to summarize and investigate the potential of ultrahigh-field imaging in neurodegenerative research. Finally, the distance between appearance of the different subfields is given in millimetres, although lengths of the MTL and hippocampus differ between individuals and might be affected by disease. It is unclear how this affects subfield measurements in the current protocol; a limitation true for all current segmentation protocols. Although a potential solution could be to provide relative distances rather than absolute distances between subregions, based on the total length of the MTL; this would further complicate the protocol. Additionally, the relative distance between subfields is not necessarily similar between subjects nor are they similarly affected by disease. Using this measure would therefore inherently also induce a measurement error.

The current protocol is not meant to replace the protocol of the harmonization effort for hippocampal subfields ([www.hippocampalsubfields.com](http://www.hippocampalsubfields.com)) or hamper the progress of this collaborative effort of many groups in various disciplines aiming to harmonize all the different protocols for hippocampal and parahippocampal subregions (Wisse et al. 2016b). Due to the iterative and thorough nature of the harmonization effort, the timeline for development of protocols for parahippocampal subregions and 7T images are further down the road and the current protocol is therefore meant to facilitate the segmentation of MTL regions, and especially parahippocampal subregions, for centers with a 7T scanner in the meantime.

In summary, we present a protocol to delineate medial temporal lobe structures as well as hippocampal subfields and provide evidence that it can be reliably applied. The inclusion of the most recent anatomical literature guiding the detailed subdivision of MTL regions and hippocampal subfields will make this an especially useful protocol for the investigation of the functional role of subregions in the MTL using fMRI, as well as research on the effect of exercise on MTL subregions and their differential relation with depression, autism, aging and neurodegenerative diseases.

## 6.5. CONTRIBUTIONS

The protocol was developed in a collaborative effort with members of the Institute of Cognitive Neurology and Dementia Research, Magdeburg (IKND), Penn Image Computing and Science Laboratory at the University of Pennsylvania, Philadelphia, and Allen Institute for Brain Science, Seattle; i.e. David Berron, Anne Hochkeppler, Laura Wisse, John Pluta, Song-Lin Ding, Anne Maass, Anica Luther, Long Xie, David Wolk, Sandhitsu Das, Thomas Wolbers, Paul Yushkevich, and Emrah Düzel. Most notably, David Berron, Laura Wisse, Anne Hochkeppler, John Pluta and I developed the initial set of rules. Song-Lin Ding provided neuroanatomical insights and reviewed segmentations. Anne Hochkeppler and Anica Luther segmented all datasets for reliability testing. David Berron and Laura Wisse did the reliability analyses. I produced all schematics and figures except for the Bland-Altman plots (by Laura Wisse). David Berron, Laura Wisse and I refined the rules and wrote the manuscript. All authors edited the manuscript.

# 7. GENERAL DISCUSSION

## 7.1. SUMMARY

The work presented in this thesis adds to the understanding of human memory, and especially of pattern completion and separation. In chapter 2, I have developed a novel recognition memory paradigm (named Memory Image Completion – MIC) particularly targeting pattern completion processes by manipulating stimulus completeness. Simultaneously, I have identified age-related recognition memory deficits suggesting a bias towards- but also a deficit in pattern completion. In chapter 3, I have replicated the findings of chapter 2, and eliminated perceptual confounds in memory performance using concurrent eye-tracking. The observed viewing patterns during encoding and retrieval could not account for the recognition memory differences across conditions and age groups, lending more validity to the task as a tool to assess pattern completion. In chapter 4, results of a collaborative case study have presented direct evidence that the hippocampus is differentially involved in pattern separation and completion. More precisely, a patient with selective bilateral DG lesions presented with memory performance indicative of deficient pattern separation, intact pattern completion, and a bias towards the latter. This was some of the first evidence directly linking the DG to pattern separation, while simultaneously excluding major contributions to pattern completion. Instead, a lesioned DG may send CA3 into overdrive promoting increased pattern completion. In chapter 5, I tried to further tackle hippocampal subfield involvement in the MIC, however, unfortunately, pattern similarity analyses remained inconclusive. Nevertheless, activity associated with pattern completion seemed to involve the STS, which indicates that a successfully retrieved pattern is reinstated there. Interestingly, some prominent age effects could be identified. Although the hippocampus was involved in more general retrieval in young but not older adults, overall older adults showed hyperactivity in the hippocampus and specifically CA3 suggesting that the hippocampal neural circuit does change with age. Additionally, generally reduced PhC-activity alongside a specific reduction during novelty processing revealed another affected site in aging. In chapter 6, a new segmentation protocol was developed in cooperation with other groups to enable accurate analyses on MTL regions including PhC, PrC, ErC and all hippocampal subfields, because the neuroanatomical literature has advanced in recent years and it is important to ensure that MR research is based on the appropriate anatomy.

Altogether, the findings presented in this thesis contribute to the literature on pattern separation and completion, and provide a new reliable means of assessing the two processes with a focus on the latter. Further on, the results support and advance existing theories of memory and aging, but also question some of its more specific predictions in MR research. Finally, this work has improved the current methodology in identifying and segmenting MTL subregions so that they are more consistent with actual neuroanatomy.

## 7.2. PATTERN COMPLETION TARGETED BY A RECOGNITION MEMORY TASK

Pattern separation and completion have received wide attention in neuroscience to foster the understanding of how the brain encodes and retrieves memories (for review, see Yassa and Stark 2011). Over recent years, the focus on humans has increased after computational models (Marr 1971; McClelland 1994; Treves and Rolls 1994; O'Reilly et al. 1998; Hasselmo and McClelland 1999) and rodent studies (Nakazawa et al. 2002; Guzowski et al. 2004; Vazdarjanova and Guzowski 2004; Leutgeb et al. 2004, 2007; Leutgeb and Leutgeb 2007) had implicated some underlying principles and brain regions (see Introduction for details 1.3). By virtue of behavioural tasks, human research has tried to approximate these processes. However, it should be clear that they are defined based on neural computations, and research in humans investigates the assumed behavioural outcome of these computations. The main emphasis has, so far, been put on pattern separation in the frequent use of the MST – a task employing very similar stimuli amongst novel and repeated stimuli with concurrent fMRI drawing inferences about pattern separation and completion (Kirwan and Stark 2007; Bakker et al. 2008; Lacy et al. 2011; etc.). While it was useful to induce memory interference, the use of highly similar stimuli during encoding selectively favoured pattern separation because the similar traces needed to be made more distinct. It is likely that pattern completion was merely a secondary finding in all these studies without being explicitly manipulated. Hunsaker and Kesner (2013 p. 40) have also picked up on this biased design, and suggested that parts of an original cue may engage pattern completion more independently than similar versions of it. Additionally, they have also remarked that though both processes are likely contributing to both, pattern separation may be more involved in encoding, while pattern completion may be more involved in retrieval. Thus, in order to increase the impact of either pattern separation or completion, studies should focus on the corresponding phase of memory processing respectively. This idea receives additional support by an eye-tracking study showing that trials in the MST interpreted to involve pattern completion (false alarms to lure, i.e. judging a similar item as old) had received fewer fixations during encoding suggesting additional involvement of pattern separation processes (Molitor et al. 2014).

Accommodating the above suggestions, I have developed the Memory Image Completion (MIC) task (see chapter 2) showing that gradually less complete versions of a learned stimulus reduce accurate recognition memory, which is interpreted to reflect an increase in pattern completion demands. To reiterate, I have used a learning criterion to prevent insufficient encoding and partial stimuli during a retrieval task to adequately trigger retrieval.

A recent paper has systematically reviewed all existing behavioural paradigms designed to test pattern separation and completion in humans including the MIC (Liu et al. 2016). The authors defined several parameters assessing task validity on the basis of two reviews (Hunsaker and Kesner 2013; Deuker et al. 2014) all of which were fulfilled in the MIC (relating to criteria for pattern completion): (1) tested during a retrieval task, (2) used partial rather than degraded cues, (3) parametrically altered degree of degradation,

and (4) accounted for confounding factors. This is a promising assessment of the MIC lending it more credibility as a suitable paradigm to assess pattern completion.

Further on, I could show in chapter 3 that the MIC can be used reliably by replicating the results from chapter 2, which indicates that they were not merely incidental findings. Additionally, the obtained eye-movement patterns showed no encoding differences, and neither could they account for the observed performance profile during retrieval. Specifically, I did not observe differential viewing patterns for new as opposed to learned stimuli, even though the literature suggests that previously seen items should receive higher fixation numbers and durations (Hannula et al. 2012). However, there are other studies that did not observe fixation differences between encoding and retrieval (Foulsham and Underwood 2008) in line with the lack of an effect here. In turn, several studies have indicated that increased fixation rates during encoding are associated with better memory representations and performance (Loftus 1972; Pertzov et al. 2009) and could thus influence retrieval (Molitor et al. 2014). As I did not observe different fixation numbers during the learning phase, stimuli were allegedly equally well encoded leading me to assume the retrieval phase was largely unaffected by encoding differences. However, the results from the imaging study in chapter 5 have cast a slightly different light onto these findings. Due to not presenting new stimuli in full, performance for their partial versions drastically dropped. Consequently, I inferred that even in the study phase stimuli were still encoded, which had become too hard without seeing the full stimuli. Analyses across sessions revealed then also that learning was taking place, because participants improved over the course of the task. This, however, implies that in the original version of the task, stimuli were still encoded during the retrieval phase which might explain the young adults' good performance in recognizing new partial stimuli. These results show that even when cautiously designing tasks to fit either pattern separation or completion demands, an interplay between the two processes cannot be excluded, and is in fact likely to occur. Consequentially, I suggest to always study them together. That is, as long as there is no underlying neural proof of the processes involved, their potential contributions to performance should always be discussed in concert. Ideally, both encoding and retrieval phases should be included and analysed in each task to have a better means of discussing their respective impact and likely interaction.

### 7.3. RECOGNITION MEMORY DIFFERENCES IN AGING RELATED TO PATTERN COMPLETION

Neurodegeneration in aging with the corresponding circuit changes in the hippocampus have been indicated to lead to an imbalance in function (Wilson et al. 2006). Specifically, because the perforant path degenerates, DG and CA3 receive less sensory input. While this leads to reduced activity in DG, CA3 is hyperactive due to its intact recurrent connections alongside reduced inhibition through less cholinergic modulation (see Introduction 1.4.1 for details). Given that DG is hypothesized to be mainly involved in pattern separation, and CA3 in pattern completion (Rolls and Kesner 2016), the consequences of these age-related changes should manifest in a deficit in pattern separation and a bias towards pattern completion (for a recent review, see Leal

and Yassa 2015). With their specific contributions to encoding and retrieval, the system is consequently assumed to favour the retrieval of already stored memories to the detriment of encoding new events (Wilson et al. 2006). Again, using the MST, there is some evidence that this is in fact the case in older adults, because they failed to correctly identify lures as similar and thought that they were old instead, which was interpreted as a deficit in pattern separation (Toner et al. 2009; Holden et al. 2013). Based on the same behavioural measure but with additional correlated CA3/DG hyperactivity, two studies suggested that this implicates a pattern completion bias (Yassa et al. 2011a, 2011b). Due to the reasons given above (see 7.2, but also Aim 1.5), these paradigms are, however, less suited to identify pattern completion. In contrast, using the MIC, I could identify a deficit in pattern completion with a bias towards it in chapters 2, 3 and 5. Crucially, these are two distinct findings with one being consistent with the models and previous findings, whereas the other has received less attention. As discussed above, the bias toward pattern completion has been proposed to occur with age resulting in preferred retrieval over new encoding. In line with this theory, the data presented in this thesis show that older adults were specifically impaired in identifying new stimuli, and frequently resorted to selecting a familiar response. Even when they made errors for learned items, they more often identified them as a different learned item than stating they had not seen it before. Thus, older participants were biased to retrieve previously learned stimuli and did not encode new ones. On the other hand, however, I also observed a deficit in pattern completion, because older adults had difficulties identifying learned items. This has not been a direct prediction of the models described above, however, a process imbalance favouring pattern completion does not necessarily mean that pattern completion is still functioning correctly. After all, CA3 also gets less sensory input implying that it has fewer information to work with. Findings in older adults showing deficient navigation to a target in an environment stripped of its original landmarks gives some additional evidence to an age-related deficit in pattern completion (Paleja and Spaniol 2013).

The eye-movement patterns from chapter 3 lend more support to these findings. More precisely, age groups did not differ in eye-movement behaviour except for one condition which might be explained by a general reduction in processing speed (see Discussion of chapter 3.4 for details). Importantly, as there were no differences viewing learned or new images, the specific impairment of older adults cannot be explained by their perceptual behaviour. These results support the idea that the observed performance differences result from mnemonic rather than perceptual impairments.

One major finding, which has been proposed in the literature regarding the age-related imbalance between pattern separation and completion, has been observed in chapter 5. Intriguingly, bilateral CA3 showed condition-independent hyperactivity in older adults. A previous study has reported something similar before; however, they directly compared age groups' activity levels within hippocampal subfields in a very specific contrast without checking other brain areas, and importantly, without differentiating between CA3 and the DG (Yassa et al. 2011a). Thus, the study in chapter 5 has been

the first to show CA3-hyperactivity in humans so clearly, given that it was not based on a task-specific contrast and stood out in a whole-scan comparison.

#### 7.4. IMPLICATIONS FOR GENERAL AGE-RELATED MEMORY PERFORMANCE

Considering that human research on pattern separation and completion, and consequently the MIC, are based on assumptions about behavioural outcomes of neural computations, it is possible that the underlying neural mechanisms are not the ones hypothesized. Therefore, I will discuss other more psychological concepts which do also fit the data but may rely on other mechanisms.

First of all, the dual process model literature differentiates between familiarity and recollection (Yonelinas 2002). Familiarity entails a sense of vaguely knowing that the event has happened, while recollection retrieves specific details of the event (see Introduction 1.2 for details). The MIC dwells more on recollection as opposed to familiarity. Importantly, in this task participants have to specifically identify the stimuli rather than just indicate whether they are old, new or similar as has been common practice in most of the studies discussed so far (e.g. Stark et al. 2015). Even studies specifically looking at differences between recollection and familiarity do use this unspecific judgement, however, with additional remember/know (e.g. Kim and Yassa 2013) or confidence indication (“sure old”, “sure new”; e.g. Koen and Rugg 2016). Nevertheless, while I did not design the experiment to make specific claims about familiarity and recollection, the findings indicate that recollection is impaired with age. This deficit manifested in older adults identifying learned stimuli less accurately than young adults, which has been consistently reported in the literature (for review, see Koen and Yonelinas 2014).

In a related field, greater impairment in free recall as opposed to recognition memory was reported in aging (Danckert and Craik 2013; Luo and Craik 2008). While the MIC does not require completely free recall, the cues prompting retrieval are severely diminished. For both young and older adults, performance gradually declined with decreasing stimulus completeness. However, older adults were more impaired suggesting that recognition memory in aging declines more with decreasing cue-availability supporting the findings for free recall. The paradigm may therefore bridge a gap between completely self-initiated recall to identical repetitions cueing retrieval (Craik 1983; Luo and Craik 2008).

Further on, it stands to reason that novelty detection contributed to performance, because older adults were impaired in identifying something new. However, as discussed in chapter 2, the linear performance decline speaks against a binary match/mismatch (old/new) mechanism in the hippocampus (Kumaran and Maguire 2009), but more likely conveys the degree of familiarity. Interestingly though, novelty processing seemed to be involved in some way as was observed with a novelty-related contrast in fMRI in chapter 5, where the PhC was differentially activated in young adults only. While this was in line with a paper suggesting specific PhC involvement in scene novelty (Howard et al. 2011), age-related changes in PhC associated with novelty have not yet been reported to my knowledge. Given, however, that PhC-activity was



reduced in older adults across the whole task, the novelty finding may be less specific. In addition, performance for new stimuli was notoriously low in both age groups, suggesting that novelty had not correctly been detected, which undermines the novelty account even more.

Finally, the results are in line with frequent reports of increased false alarm rates in older adults (Schacter et al. 1997). Although false alarms in the MIC are slightly different to other paradigms, the findings can inform the debate on the reason for increased false alarm rates with more detail. While usually a false alarm indicates an 'old' response to a new item, false alarms in the MIC were more specific in that a particular old stimulus was chosen for an answer. Thus, false alarms in the MIC do not reflect a vague sense of familiarity, but rather incorrect recollection. Furthermore, there are two categories of false alarms in the MIC: false alarms to new items and false alarms to old items. That is, while a wrong response for a learned item would just indicate a 'miss' in standard paradigms, in the MIC that is only true when 'new' is selected as a response. However, the data of chapters 2, 3 and 5 show that older adults frequently do not 'miss' a learned stimulus, but rather incorrectly retrieve another one. This provides evidence that there is not a global criterion shift in aging like an overall response bias to say 'new' (as suggested by Schacter et al. 1997), but a rather specific impairment in memory retrieval. The fMRI data from chapter 5 further support this notion, because retrieval-related hippocampal activity was not observed in older adults.

### 7.5. BRAIN REGIONS ASSOCIATED WITH THE MIC

As presented in numerous places in this thesis, computational models of hippocampal function predict a very specific functional differentiation between DG and CA3 with the former serving pattern separation and the latter serving pattern completion (for a recent review, see Rolls and Kesner 2016). This has recently been confirmed in rodents (Neunuebel and Knierim 2014). Additionally, CA1 may function as an integration hub between sensory information from ErC and pre-processed DG and CA3 output (Lee et al. 2004). While there is plenty of evidence suggesting a combined contribution of DG and CA3 to mnemonic discrimination interpreted as pattern separation (for review, see Yassa and Stark 2011), there is yet only one human fMRI study suggesting involvement of DG only in pattern separation (Berron et al. 2016). Thus, the findings presented in chapter 4 are unique in lending direct support for the distinct neural mechanisms associated with the MST and the MIC. Specifically, a patient with bilateral DG lesions was drastically impaired in discriminating similar stimuli, suggesting a severe deficit in pattern separation (Baker et al. 2016). Simultaneously, he was *not* impaired in correctly retrieving learned stimuli from partial cues, implying that pattern completion was intact. Additionally, he presented with a substantial tendency to pick familiar responses indicating a bias toward pattern completion. These results underpin the assumption that DG is crucial for pattern separation. In addition, a bias toward pattern completion likely results from deficient DG output as a basis for further processing in CA3. Simultaneously, DG is not necessary for successful pattern completion, leaving room for CA3 playing its part. To identify the latter, I designed an fMRI study which is

described in chapter 5. Derived from the proposed computations of orthogonalizing and equalizing memory representations, a pattern similarity analysis comparing representations across voxels within specific hippocampal subfields seemed most promising and best suited to identify pattern separation and completion. However, the findings were inconclusive. Despite strong univariate effects, I could not detect differential patterns in hippocampal subfields. While some research succeeded in doing so (e.g. Kyle et al. 2015a; Stokes et al. 2015), there are studies that also failed to identify hippocampal contributions with pattern similarity analyses (LaRocque et al. 2013). Nevertheless, as already discussed above, the hippocampus was involved in unspecific retrieval in young adults, and general CA3-hyperactivity was observed in older adults in line with the model of cognitive aging (Wilson et al. 2006). However, I identified cortex involvement in the STS which was more specific to memory comparisons likely conveying contributions of pattern completion. That is, when contrasting conditions in which I assumed successful pattern completion (e.g. when a partial learned image was correctly identified) with conditions where I did not (when correctly identifying a new image), STS showed higher activity for the latter. This is based on a repetition suppression account, assuming that the repeated presentation of a stimulus results in less activation. Consequently, if a stimulus is completed to its original form it should follow the same characteristics as a repetition (see Yassa and Stark 2011 for details). Importantly, using this rationale, the results of a process (the completed representation) can be identified rather than the process itself. Crucially, models of pattern separation and completion predict that a successfully completed representation is subsequently reinstated in cortical regions (Treves and Rolls 1994; McClelland et al. 1995) with a recent suggestion that this could in fact take place in STS (Kesner and Rolls 2015). Given that STS seemed to reflect the behavioural response (thus, a completed representation) it is likely that some process has preceded its activation, potentially indicating cortical reinstatement following successful pattern completion. As indicated before, 'active' pattern completion could not be identified, neither with univariate nor multivariate analyses. However, following the reasoning about repetition suppression above, trying to identify stimulus-specific similarity may not be the best approach to study this, as it assumes similar representations between stimuli which reflect the result of a process rather than the process itself. It may thus be worthwhile to employ a different approach trying to identify processes rather than stimulus-specific representations.

Lastly, CA3-hyperactivity that was observed in older adults is a promising finding in favour of aging models (Wilson et al. 2006) given its global occurrence. However, it should be noted that by using any of the existing segmentation protocols, this would not have been such a clear result. That is, big parts of where hyperactivity in older adults was observed would have been identified as DG when applying the former rules. Critically, most protocols have so far either not differentiated between CA3 and DG (e.g. Frisoni et al. 2015), or underestimated CA3 mainly because of insufficient resolution or contrast (e.g. Wisse et al. 2012; Winterburn et al. 2013). Although I have already employed a different border in chapter 5 by dividing CA3 and DG along the

endfolial pathway, it became clear that there were more discrepancies between existing protocols and the newest neuroanatomical findings (for a recent atlas, see Ding et al. 2016). Therefore, in chapter 6 a new extensive segmentation protocol for MTL subregions was developed in a collaborative effort, in order to define rules which better resemble neuroanatomy. Additionally, regions in the parahippocampal gyrus were another big feature of the protocol due to their susceptibility to great individual variations, which have now been incorporated in the rules. Given that PhC was another region differently involved in aging in chapter 5 clearly defining its border is necessary in order to make claims about its specific contribution especially as it posteriorly stretches into visual areas which likely serve different functions.

## 7.6. OUTLOOK AND FUTURE PERSPECTIVES

The results in this thesis have shown that a behavioural recognition memory task is suited to address questions related to pattern completion, and can be reliably used. Additionally, some neural mechanisms have been identified; especially chapter 4 showed that DG played a crucial role in a task approximating pattern separation, whereas it did not seem to be necessary for performance related to pattern completion. However, the exact underlying neural basis still needs to be determined, given that the imaging results presented in chapter 5 were inconclusive regarding specific hippocampal subfield involvement in solving the task. However, the univariate analyses did identify some unspecific contribution of the whole hippocampus to retrieval, alongside unequivocal age-effects mainly in CA3. Further on, the STS was associated with contrasts suggesting cortical reinstatement after successful pattern completion. It may therefore be worthwhile to investigate representational similarity in STS identical to the analysis done in hippocampal subfields, to get a better understanding if stimulus-specific representations are reinstated there or not. In contrast, a process analysis of hippocampal subfields may prove more promising than representational similarity. This could possibly be done by calculating multi-voxel pattern correlations across entire conditions and not between single trials, that is, a similar behaviour-dependent GLM as was used for univariate analyses could serve as a basis for a process pattern similarity analysis. On the other hand, pattern similarity analyses might not be sensitive enough overall to identify differences between voxel patterns. Alternatively, a pattern classification approach may be used instead, given that it has been applied successfully in classifying between stimulus representations in CA1, CA3 (Bonnici et al. 2012a) and DG (Berron et al. 2016). Critically, however, the development of standard tools especially tuned for 7T-fMRI analysis is indispensable. As shortly outlined above (see chapter 5.4), common software often fails in ordinary pre-processing like coregistration and normalization. In addition, whole-scan multiple comparisons corrections do usually not spare any significant results, because 7T scans usually contain considerably more voxels and are not or only slightly smoothed compared to 3T scans. This suggests that potentially other threshold-criteria may have to be applied, but which are still complying with good scientific conduct in drawing statistical inferences.

Apart from that, the behavioural results from all experimental chapters show some very specific age differences. Although merely functioning as a control group, the data of participants in chapter 4 is also quite informative. Participants were middle aged (around 50 years old), thus, in a rarely studied age group, and showed a starting shift in their bias curve. Albeit not being significantly different from 0, their bias curve had a shape more similar to older adults than to young adults (from chapters 2 and 3). Therefore, it is interesting to look at behavioural changes using the MIC across the whole life span, to see when and how the performance decline progresses. Developing a database with age-typical population responses can consequently inform about deviations from the norm potentially revealing pathological behaviour. Importantly, such a tool could inform about two distinct measures at the same time: (1) impaired recognition memory associated with a deficit in pattern completion, and (2) a differential tendency to choose familiar responses indicating a bias towards pattern completion. Results from chapter 4 showed that those do not necessarily go hand in hand. Furthermore, it could be useful to generate several versions of the task to be able to test participants at different time points without learning transfer. This may inform about individual performance decline over an extended period of time.

Given that I have found reduced PhC-activity in older adults in chapter 5, which is a region suggested to be more sensitive in differentiating between healthy and pathological aging (Echávarri et al. 2011), it is also desirable to identify behaviour and concurrent brain changes in populations with MCI, early or late AD. Importantly, the segmentation protocol presented in chapter 6 can advance the structural and functional specificity of future studies by providing a means to better fit accurate neuroanatomy and identify regions affected by pathological aging.

Finally, terminology in the studies of pattern separation and completion in humans is sometimes used interchangeably and can be confusing in the context of these processes. For example, pattern separation is also referred to as discrimination (De Shetler and Rissman 2016), lure discrimination (e.g. Reagh and Yassa 2014b) or mnemonic discrimination (e.g. Bennett and Stark 2016), whereas pattern completion is sometimes substituted by generalization (De Shetler and Rissman 2016) or reinstatement (e.g. Staresina et al. 2012). This practice is inadequate and can be gravely misleading given that there is a huge variety of research where concepts sometimes have similar or even the same names (e.g. generalization in the context of fear conditioning; Xu and Südhof 2013), are unspecific (e.g. perceptual discrimination independent of memory; Lee et al. 2005b), or – in the worst case – are very close yet different from the process in question (e.g. cortical reinstatement following pattern completion; Treves and Rolls 1994; Rolls 2016). While most of the terms refer to some sort of behavioural epiphenomenon, researchers should make very clear if they study the underlying neural computation or the behavioural outcome, and should make the connections between the two very clear rather than using the terminology interchangeably.

## 8. REFERENCES

- Aggleton JP, Brown MW. 2006. Interleaving brain systems for episodic and recognition memory. *Trends Cogn Sci* 10: 455–63.
- Ally BA, Hussey EP, Ko PC, Molitor RJ. 2013. Pattern separation and pattern completion in Alzheimer's disease: Evidence of rapid forgetting in amnesic mild cognitive impairment. *Hippocampus* 23: 1246–58.
- Aly M, Ranganath C, Yonelinas AP. 2013. Detecting changes in scenes: the hippocampus is critical for strength-based perception. *Neuron* 78: 1127–37.
- Aly M, Turk-Browne NB. 2015. Attention Stabilizes Representations in the Human Hippocampus. *Cereb Cortex* 1–14.
- Amaral DG. 1978. A golgi study of cell types in the hilar region of the hippocampus in the rat. *J Comp Neurol* 182: 851–914.
- Amaral DG, Witter MP. 1989. The three-dimensional organization of the hippocampal formation: A review of anatomical data. *Neuroscience* 31: 571–591.
- Apostolova LG, Zarow C, Biado K, Hurtz S, Boccardi M, Somme J, Honarpisheh H, Blanken AE, Brook J, Tung S, et al. 2015. Relationship between hippocampal atrophy and neuropathology markers: A 7T MRI validation study of the EADC-ADNI Harmonized Hippocampal Segmentation Protocol. *Alzheimer's Dement* 11: 139–150.
- Avants BB, Tustison NJ, Song G, Cook PA, Klein A, Gee JC. 2011. A reproducible evaluation of ANTs similarity metric performance in brain image registration. *Neuroimage* 54: 2033–2044.
- Azab M, Stark SM, Stark CEL. 2014. Contributions of human hippocampal subfields to spatial and temporal pattern separation. *Hippocampus* 24: 293–302.
- Baker S, Vieweg P, Gao F, Gilboa A, Wolbers T, Black SE, Rosenbaum RS. 2016. The Human Dentate Gyrus Plays a Necessary Role in Discriminating New Memories. *Curr Biol* 26: 2629–2634.
- Bakker A, Kirwan CB, Miller MI, Stark CEL. 2008. Pattern separation in the human hippocampal CA3 and dentate gyrus. *Science* 319: 1640–1642.
- Bakker A, Krauss GL, Albert MS, Speck CL, Jones LR, Stark CEL, Yassa MA, Bassett SS, Shelton AL, Gallagher M. 2012. Reduction of hippocampal hyperactivity improves cognition in amnesic mild cognitive impairment. *Neuron* 74: 467–74.
- Bennett IJ, Stark CEL. 2016. Mnemonic discrimination relates to perforant path integrity: An ultra-high resolution diffusion tensor imaging study. *Neurobiol Learn Mem* 129: 107–112.
- Bernasconi N, Bernasconi A, Caramanos Z, Antel SB, Andermann F, Arnold DL. 2003. Mesial temporal damage in temporal lobe epilepsy: a volumetric MRI study of the hippocampus, amygdala and parahippocampal region. *Brain* 126: 462–9.
- Berron D, Schütze H, Maass A, Cardenas-Blanco A, Kuijff HJ, Kumaran D, Düzel E. 2016. Strong Evidence for Pattern Separation in Human Dentate Gyrus. *J Neurosci* 36: 7569–7579.
- Berron D, Vieweg P, Hochkeppler A, Pluta JBB, Maass A, Luther A, Das SRR, Wolk DAA, Wolbers T, Yushkevich PAA, et al. 2017. A protocol for manual segmentation of medial temporal lobe subregions in 7 Tesla MRI. *NeuroImage Clin* 15: 466–482.
- Bonnici HM, Chadwick MJ, Kumaran D, Hassabis D, Weiskopf N, Maguire EA. 2012a. Multi-voxel pattern analysis in human hippocampal subfields. *Front Hum Neurosci* 6: 290.
- Bonnici HM, Kumaran D, Chadwick MJ, Weiskopf N, Hassabis D, Maguire EA. 2012b. Decoding representations of scenes in the medial temporal lobes. *Hippocampus* 22: 1143–53.
- Boutet C, Chupin M, Lehericy S, Marrakchi-Kacem L, Epelbaum S, Poupon C, Wiggins CJ, Vignaud A, Hasboun D, Defontaine B, et al. 2014. Detection of volume loss in hippocampal layers in Alzheimer's disease using 7 T MRI: a feasibility study. *NeuroImage Clin* 5: 341–8.
- Braak H, Braak E. 1991. Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol* 82: 239–59.
- Brainerd CJ, Reyna VF. 2001. Fuzzy-trace theory: dual processes in memory, reasoning, and cognitive neuroscience. In *Advances in Child Development and Behavior*, Vol. 28 of, pp. 41–100.
- Brainerd CJ, Reyna VF. 2002. Fuzzy-Trace Theory and False Memory. *Curr Dir Psychol Sci* 11: 164–169.
- Brenner D, Stirnberg R, Pracht ED, Stöcker T. 2014. Two-dimensional accelerated MP-RAGE imaging with flexible linear reordering. *Magn Reson Mater Physics, Biol Med* 27: 455–462.
- Bronen RA, Cheung G. 1991. Relationship of hippocampus and amygdala to coronal MRI landmarks. *Magn Reson Imaging* 9: 449–457.

- Burgmans S, van Boxtel MPJ, van den Berg KEM, Gronenschild EHBM, Jacobs HIL, Jolles J, Uylings HBM. 2011. The posterior parahippocampal gyrus is preferentially affected in age-related memory decline. *Neurobiol Aging* 32: 1572–1578.
- Burke SN, Barnes C a. 2006. Neural plasticity in the ageing brain. *Nat Rev Neurosci* 7: 30–40.
- Burke SN, Wallace JL, Nematollahi S, Uprety AR, Barnes C a. 2010. Pattern separation deficits may contribute to age-associated recognition impairments. *Behav Neurosci* 124: 559–73.
- Colgin LL, Moser EI, Moser M-B. 2008. Understanding memory through hippocampal remapping. *Trends Neurosci* 31: 469–77.
- Copara MS, Hassan AS, Kyle CT, Libby LA, Ranganath C, Ekstrom AD. 2014. Complementary Roles of Human Hippocampal Subregions during Retrieval of Spatiotemporal Context. *J Neurosci* 34: 6834–6842.
- Corwin J, Bylsma FW. 1993. Psychological examination of traumatic encephalopathy. *Clin Neuropsychol* 7: 3–21.
- Craik FIM. 1983. On the transfer of information from temporary to permanent memory. *Philos Trans R Soc Lond B Biol Sci* 302: 341–359.
- Crawford JR, Garthwaite PH. 2002. Investigation of the single case in neuropsychology: confidence limits on the abnormality of test scores and test score differences. *Neuropsychologia* 40: 1196–1208.
- Crawford JR, Garthwaite PH. 2005. Testing for Suspected Impairments and Dissociations in Single-Case Studies in Neuropsychology: Evaluation of Alternatives Using Monte Carlo Simulations and Revised Tests for Dissociations. *Neuropsychology* 19: 318–331.
- Crawford JR, Howell DC. 1998. Comparing an Individual's Test Score Against Norms Derived from Small Samples. *Clin Neuropsychol (Neuropsychology, Dev Cogn Sect D)* 12: 482–486.
- Danckert SL, Craik FIM. 2013. Does aging affect recall more than recognition memory? *Psychol Aging* 28: 902–9.
- Das SR, Pluta JB, Mancuso L, Kliot D, Yushkevich PA, Wolk DA. 2015. Anterior and posterior MTL networks in aging and MCI. *Neurobiol Aging* 36: 141–50.
- Daugherty AM, Bender AR, Raz N, Ofen N. 2015. Age Differences in Hippocampal Subfield Volumes from Childhood to Late Adulthood. *Hippocampus*.
- Davis T, LaRocque KF, Mumford JA, Norman KA, Wagner AD, Poldrack RA. 2014. What do differences between multi-voxel and univariate analysis mean? How subject-, voxel-, and trial-level variance impact fMRI analysis. *Neuroimage* 97: 271–83.
- de Flores R, Joie R, Landeau B, Perrotin A, Mézenge F, de Sayette V, Eustache F, Desgranges B, Chételat G. 2015a. Effects of age and Alzheimer's disease on hippocampal subfields: comparison between manual and FreeSurfer volumetry. 36: 463–74.
- de Flores R, La Joie R, Chételat G. 2015b. Structural imaging of hippocampal subfields in healthy aging and Alzheimer's disease. *Neuroscience* 309: 29–50.
- De Shetler NG, Rissman J. 2016. Dissociable profiles of generalization/discrimination in the human hippocampus during associative retrieval. *Hippocampus* 121: 1–22.
- Derdikman D, Knierim JJ. 2014. *Space, Time and Memory in the Hippocampal Formation*. eds. D. Derdikman and J.J. Knierim. Springer Vienna, Vienna.
- Deuker L, Doeller CF, Fell J, Axmacher N. 2014. Human neuroimaging studies on the hippocampal CA3 region - integrating evidence for pattern separation and completion. *Front Cell Neurosci* 8: 64.
- Diana RA, Yonelinas AP, Ranganath C. 2007. Imaging recollection and familiarity in the medial temporal lobe: a three-component model. *Trends Cogn Sci* 11: 379–386.
- Dice LR. 1945. Measures of the Amount of Ecologic Association Between Species. *Ecology* 26: 297–302.
- Dickerson BC, Salat DH, Bates JF, Atiya M, Killiany RJ, Greve DN, Dale AM, Stern CE, Blacker D, Albert MS, et al. 2004. Medial temporal lobe function and structure in mild cognitive impairment. 56: 27–35.
- Diersch N. 2013. Action prediction in the aging mind. Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig.
- Ding S. 2015. Detailed segmentation of human hippocampal and subicular subfields using a combined approach. *Neurosci Commun* 1: 1–9.
- Ding S-L. 2013. Comparative anatomy of the prosubiculum, subiculum, presubiculum, postsubiculum, and parasubiculum in human, monkey, and rodent. *J Comp Neurol* 521: 4145–4162.
- Ding S-L, Royall JJ, Sunkin SM, Ng L, Facer BAC, Lesnar P, Guillozet-Bongaarts A, McMurray B, Szafer A, Dolbeare TA, et al. 2016. Comprehensive cellular-resolution atlas of the adult human brain. *J Comp Neurol* 524: 3127–3481.

- Ding S-L, Van Hoesen GW. 2010. Borders, extent, and topography of human perirhinal cortex as revealed using multiple modern neuroanatomical and pathological markers. *Hum Brain Mapp* 31: 1359–1379.
- Ding S-L, Van Hoesen GW. 2015. Organization and detailed parcellation of human hippocampal head and body regions based on a combined analysis of cyto- and chemo-architecture. *J Comp Neurol* 2253: 2233–2253.
- Ding S-L, Van Hoesen GW, Cassell MD, Poremba A. 2009. Parcellation of human temporal polar cortex: A combined analysis of multiple cytoarchitectonic, chemoarchitectonic, and pathological markers. *J Comp Neurol* 514: 595–623.
- Dowiasch S, Marx S, Einhäuser W, Bremmer F. 2015. Effects of aging on eye movements in the real world. *Front Hum Neurosci* 9: 46.
- Du A-T, Schuff N, Kramer JH, Rosen HJ, Maria G-T, Rankin K, Miller BL, Weiner MW. 2007. Different regional patterns of cortical thinning in Alzheimer's disease and frontotemporal dementia. 130: 1159–1166.
- Duarte A, Graham KS, Henson RN. 2010. Age-related changes in neural activity associated with familiarity, recollection and false recognition. *Neurobiol Aging* 31: 1814–30.
- Duncan K, Sadanand A, Davachi L. 2012. Memory's Penumbra: Episodic Memory Mnemonic Biases. *Science* 337: 485–487.
- Duncan K, Tomparay A, Davachi L. 2014. Associative Encoding and Retrieval Are Predicted by Functional Connectivity in Distinct Hippocampal Area CA1 Pathways. *J Neurosci* 34: 11188–98.
- Duvernoy HM, Cattin F, Risold P-Y, Vannson JL, Gaudron M. 2013. *The human hippocampus: functional anatomy, vascularization, and serial sections with MRI*. Springer.
- Eatough EM, Shirlcliff EA, Hanson JL, Pollak SD. 2009. Hormonal reactivity to MRI scanning in adolescents. *Psychoneuroendocrinology* 34: 1242–6.
- Echávvarri C, Aalten P, Uylings HBM, Jacobs HIL, Visser PJ, Gronenschild EHBM, Verhey FRJ, Burgmans S. 2011. Atrophy in the parahippocampal gyrus as an early biomarker of Alzheimer's disease. *Brain Struct Funct* 215: 265–71.
- Eichenbaum H, Yonelinas AP, Ranganath C. 2007. The medial temporal lobe and recognition memory. *Annu Rev Neurosci* 30: 123–152.
- Ekstrom AD, Bazih AJ, Suthana NA, Al-Hakim R, Ogura K, Zeineh MM, Burggren AC, Bookheimer SY. 2009. Advances in high-resolution imaging and computational unfolding of the human hippocampus. *Neuroimage* 47: 42–49.
- Ekstrom AD, Kahana MJ, Caplan JB, Fields TA, Isham EA, Newman EL, Fried I. 2003. Cellular networks underlying human spatial navigation. *Nature* 425: 184–188.
- Epstein RA. 2008. Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends Cogn Sci* 12: 388–396.
- Feczko E, Augustinack JC, Fischl B, Dickerson BC. 2009. An MRI-based method for measuring volume, thickness and surface area of entorhinal, perirhinal, and posterior parahippocampal cortex. *Neurobiol Aging* 30: 420–431.
- Fjell AM, Walhovd KB. 2005. Age-sensitivity of P3 in high-functioning adults. *Neurobiol Aging* 26: 1297–1299.
- Foulsham T, Underwood G. 2008. What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. *J Vis* 8: 1–17.
- Fraser MA, Shaw ME, Cherbuin N. 2015. A systematic review and meta-analysis of longitudinal hippocampal atrophy in healthy human ageing☆.
- Friedman D. 2000. Event-related brain potential investigations of memory and aging. *Biol Psychol* 54: 175–206.
- Frisoni GB, Jack CR, Bocchetta M, Bauer C, Frederiksen KS, Liu Y, Preboske G, Swihart T, Blair M, Cavedo E, et al. 2015. The EADC-ADNI harmonized protocol for manual hippocampal segmentation on magnetic resonance: Evidence of validity. *Alzheimer's Dement* 11: 111–125.
- Fyhn M, Hafting T, Treves A, Moser M-B, Moser EI. 2007. Hippocampal remapping and grid realignment in entorhinal cortex. *Nature* 446: 190–4.
- Gilbert PE, Kesner RP. 2002. The amygdala but not the hippocampus is involved in pattern separation based on reward value. *Neurobiol Learn Mem* 77: 338–53.
- Gilbert PE, Kesner RP, Lee I. 2001. Dissociating hippocampal subregions: double dissociation between dentate gyrus and CA1. *Hippocampus* 11: 626–36.
- Gillund G, Shiffrin RM. 1984. A retrieval model for both recognition and recall. *Psychol Rev* 91: 1–67.
- Gollin ES. 1960. Developmental studies of visual recognition of incomplete objects. *Percept Mot Skills* 11: 289.

- Gorbach T, Pudas S, Lundquist A, Orädd G, Josefsson M, Salami A, de Luna X, Nyberg L. 2016. Longitudinal association between hippocampus atrophy and episodic-memory decline. *Neurobiol Aging*.
- Goubran M, Rudko DA, Santyr B, Gati J, Szekeres T, Peters TM, Khan AR. 2014. In vivo normative atlas of the hippocampal subfields using multi-echo susceptibility imaging at 7 Tesla. 35: 3588–3601.
- Grady C. 2012. The cognitive neuroscience of ageing. *Nat Rev Neurosci* 13: 491–505.
- Graham KS, Barense MD, Lee ACH. 2010. Going beyond LTM in the MTL: A synthesis of neuropsychological and neuroimaging findings on the role of the medial temporal lobe in memory and perception. *Neuropsychologia* 48: 831–853.
- Grill-Spector K, Henson RN, Martin A. 2006. Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn Sci* 10: 14–23.
- Grill-Spector K, Kourtzi Z, Kanwisher N. 2001. The lateral occipital complex and its role in object recognition. *Vision Res* 41: 1409–1422.
- Gutchess AH, Welsh RC, Hedden T, Bangert A, Minear M, Liu LL, Park DC. 2005. Aging and the Neural Correlates of Successful Picture Encoding: Frontal Activations Compensate for Decreased Medial-Temporal Activity. *J Cogn Neurosci* 17: 84–96.
- Guzowski JF, Knierim JJ, Moser EI. 2004. Ensemble Dynamics of Hippocampal Regions CA3 and CA1. *Neuron* 44: 581–584.
- Hannula DE, Baym CL, Warren DE, Cohen NJ. 2012. The eyes know: eye movements as a veridical index of memory. *Psychol Sci* 23: 278–87.
- Hannula DE, Ranganath C. 2009. The eyes have it: hippocampal activity predicts expression of memory in eye movements. *Neuron* 63: 592–9.
- Hasselmo ME, McClelland JL. 1999. Neural models of memory. *Curr Opin Neurobiol* 9: 184–188.
- Hasselmo ME, Schnell E, Barkai E. 1995. Dynamics of learning and recall at excitatory recurrent synapses and cholinergic modulation in rat hippocampal region CA3. *J Neurosci* 15: 5249–62.
- Hein G, Knight RT. 2008. Superior temporal sulcus—It's my area: or is it? *J Cogn Neurosci* 20: 2125–2136.
- Hemby SE, Trojanowski JQ, Ginsberg SD. 2003. Neuron-specific age-related decreases in dopamine receptor subtype mRNAs. *J Comp Neurol* 456: 176–183.
- Henderson JM, Larson CL, Zhu DC. 2008. Full scenes produce more activation than close-up scenes and scene-diagnostic objects in parahippocampal and retrosplenial cortex: an fMRI study. *Brain Cogn* 66: 40–9.
- Holden HM, Hoebel C, Loftis K, Gilbert PE. 2012. Spatial pattern separation in cognitively normal young and older adults. *Hippocampus* 22: 1826–32.
- Holden HM, Toner CK, Pirogovsky E, Kirwan CB, Gilbert PE. 2013. Visual object pattern separation varies in older adults. *Learn Mem* 20: 358–62.
- Hollingworth A, Henderson JM. 1998. Does consistent scene context facilitate object perception? *J Exp Psychol Gen* 127: 398–415.
- Horner AJ, Bisby JA, Bush D, Lin W-J, Burgess N. 2015. Evidence for holistic episodic recollection via hippocampal pattern completion. *Nat Commun* 6: 7462.
- Horner AJ, Burgess N. 2014. Pattern completion in multielement event engrams. *Curr Biol* 24: 988–992.
- Howard LR, Kumaran D, Ólafsdóttir HF, Spiers HJ. 2011. Double dissociation between hippocampal and parahippocampal responses to object-background context and scene novelty. *J Neurosci* 31: 5253–61.
- Huang Y, Coupland NJ, Lebel RM, Carter R, Seres P, Wilman AH, Malykhin N V. 2013. Structural Changes in Hippocampal Subfields in Major Depressive Disorder: A High-Field Magnetic Resonance Imaging Study. *Biol Psychiatry* 74: 62–68.
- Hunsaker MR, Kesner RP. 2013. The operation of pattern separation and pattern completion processes associated with different attributes or domains of memory. *Neurosci Biobehav Rev* 37: 36–58.
- Iglesias JE, Augustinack JC, Nguyen K, Player CM, Player A, Wright M, Roy N, Frosch MP, McKee AC, Wald LL, et al. 2015. A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution MRI: Application to adaptive segmentation of in vivo MRI. *Neuroimage* 115: 117–137.
- In M-H, Speck O. 2012. Highly accelerated PSF-mapping for EPI distortion correction with improved fidelity. *Magn Reson Mater Physics, Biol Med* 25: 183–192.
- Insausti R, Amaral DG. 2012. Hippocampal Formation. In *The human nervous system* (eds. J.K. Mai and G. Paxinos), pp. 896–942, Elsevier Academic Press.
- Insausti R, Juottonen K, Soininen H, Insausti AM, Partanen K, Vainio P, Laakso MP, Pitkänen A. 1998. MR volumetric analysis of the human entorhinal, perirhinal, and temporopolar cortices. *Am J Neuroradiol* 19: 659–71.



- Jagust W. 2013. Vulnerable Neural Systems and the Borderland of Brain Aging and Neurodegeneration. *Neuron* 77: 219–234.
- Jung MW, McNaughton BL. 1993. Spatial selectivity of unit activity in the hippocampal granular layer. *Hippocampus* 3: 165–182.
- Kafkas A, Montaldi D. 2012. Familiarity and recollection produce distinct eye movement, pupil and medial temporal lobe responses when memory strength is matched. *Neuropsychologia* 50: 3080–3093.
- Kafkas A, Montaldi D. 2014. Two separate, but interacting, neural systems for familiarity and novelty detection: A dual-route mechanism. *Hippocampus* 0.
- Kerchner GA, Deutsch GK, Zeineh MM, Dougherty RF, Saranathan M, Rutt BK. 2012. Hippocampal CA1 apical neuropil atrophy and memory performance in Alzheimer’s disease. *Neuroimage* 63: 194–202.
- Kesner RP, Kirk RA, Yu Z, Polansky C, Musso ND. 2016. Dentate gyrus supports slope recognition memory, shades of grey-context pattern separation and recognition memory, and CA3 supports pattern completion for object memory. *Neurobiol Learn Mem* 129: 29–37.
- Kesner RP, Rolls ET. 2015. A computational theory of hippocampal function, and tests of the theory: New developments. *Neurosci Biobehav Rev* 48: 92–147.
- Kim J, Yassa MA. 2013. Assessing recollection and familiarity of similar lures in a behavioral pattern separation task. *Hippocampus* 23: 287–94.
- Kim JG, Biederman I, Lescroart MD, Hayworth KJ. 2009. Adaptation to objects in the lateral occipital complex (LOC): Shape or semantics? *Vision Res* 49: 2297–2305.
- Kirasic KC. 1991. Spatial cognition and behavior in young and elderly adults: implications for learning new environments. *Psychol Aging* 6: 10–8.
- Kirwan CB, Stark CEL. 2007. Overcoming interference: an fMRI investigation of pattern separation in the medial temporal lobe. *Learn Mem* 14: 625–633.
- Kivisaari SL, Probst A, Taylor KI. 2013. The Perirhinal, Entorhinal, and Parahippocampal Cortices and Hippocampus: An Overview of Functional Anatomy and Protocol for Their Segmentation in MR Images. In *fMRI*, pp. 239–267, Springer Berlin Heidelberg, Berlin, Heidelberg.
- Knierim JJ. 2015. The hippocampus. *Curr Biol* 25: R1116–R1121.
- Koen JD, Rugg MD. 2016. Memory Reactivation Predicts Resistance to Retroactive Interference: Evidence from Multivariate Classification and Pattern Similarity Analyses. *J Neurosci* 36: 4389–99.
- Koen JD, Yonelinas AP. 2016. Recollection, not familiarity, decreases in healthy ageing: Converging evidence from four estimation methods. 24: 75–88.
- Koen JD, Yonelinas AP. 2014. The effects of healthy aging, amnesic mild cognitive impairment, and Alzheimer’s disease on recollection and familiarity: a meta-analytic review. *Neuropsychol Rev* 24: 332–54.
- Krekelberg B, Boynton GM, van Wezel RJA. 2006. Adaptation: from single cells to BOLD signals. *Trends Neurosci* 29: 250–256.
- Kriegeskorte N, Mur M, Bandettini P. 2008. Representational similarity analysis - connecting the branches of systems neuroscience. *Front Syst Neurosci* 2: 4.
- Kumaran D, Maguire EA. 2007. Match mismatch processes underlie human hippocampal responses to associative novelty. *J Neurosci* 27: 8517–24.
- Kumaran D, Maguire EA. 2009. Novelty signals: a window into hippocampal information processing. *Trends Cogn Sci* 13: 47–54.
- Kyle CT, Smuda DN, Hassan AS, Ekstrom AD. 2015a. Roles of human hippocampal subfields in retrieval of spatial and temporal context. *Behav Brain Res* 278: 549–558.
- Kyle CT, Stokes JD, Lieberman JS, Hassan AS, Ekstrom ADD. 2015b. Successful retrieval of competing spatial environments in humans involves hippocampal pattern separation mechanisms. *Elife* 4: 160.
- La Joie R, Fouquet M, Mézenge F, Landeau B, Villain N, Mevel K, Pélerin A, Eustache F, Desgranges B, Chételat G. 2010. Differential effect of age on hippocampal subfields assessed using a new high-resolution 3T MR sequence. *Neuroimage* 53: 506–514.
- La Joie R, Perrotin A, De La Sayette V, Egret S, Dœuvre L, Belliard S, Eustache F, Desgranges B, Chételat G. 2013. Hippocampal subfield volumetry in mild cognitive impairment, Alzheimer’s disease and semantic dementia. *NeuroImage Clin* 3: 155–162.
- Lacy JW, Yassa MA, Stark SM, Muftuler LT, Stark CEL. 2011. Distinct pattern separation related transfer functions in human CA3/dentate and CA1 revealed using high-resolution fMRI and variable mnemonic similarity. *Learn Mem* 18: 15–8.

- LaRocque KF, Smith ME, Carr VA, Witthoft N, Grill-Spector K, Wagner AD. 2013. Global similarity and pattern separation in the human medial temporal lobe predict subsequent memory. *J Neurosci* 33: 5466–74.
- Leal SL, Tighe SK, Jones CK, Yassa MA. 2014. Pattern separation of emotional information in hippocampal dentate and CA3. *Hippocampus* 24: 1146–55.
- Leal SL, Yassa MA. 2015. Neurocognitive Aging and the Hippocampus across Species. *Trends Neurosci* 38: 800–812.
- Lee ACH, Buckley MJ, Pegman SJ, Spiers HJ, Scahill VL, Gaffan D, Bussey TJ, Davies RR, Kapur N, Hodges JR, et al. 2005a. Specialization in the medial temporal lobe for processing of objects and scenes. *Hippocampus* 15: 782–97.
- Lee ACH, Bussey TJ, Murray EA, Saksida LM, Epstein RA, Kapur N, Hodges JR, Graham KS. 2005b. Perceptual deficits in amnesia: challenging the medial temporal lobe “mnemonic” view. *Neuropsychologia* 43: 1–11.
- Lee H, Wang C, Deshmukh SS, Knierim JJ. 2015. Neural Population Evidence of Functional Heterogeneity along the CA3 Transverse Axis: Pattern Completion versus Pattern Separation. *Neuron* 87: 1093–1105.
- Lee I, Yoganarasimha D, Rao G, Knierim JJ. 2004. Comparison of population coherence of place cells in hippocampal subfields CA1 and CA3. *Nature* 430: 456–459.
- Lee JK, Ekstrom AD, Ghetti S. 2014. Volume of hippocampal subfields and episodic memory in childhood and adolescence. *Neuroimage* 94: 162–171.
- Lehrl S, Triebig G, Fischer B. 1995. Multiple choice vocabulary test MWT as a valid and short test to estimate premorbid intelligence. *Acta Neurol Scand* 91: 335–345.
- Lerner Y, Hendler T, Malach R. 2002. Object-completion Effects in the Human Lateral Occipital Complex. *Cereb Cortex* 12: 163–177.
- Leutgeb JK, Leutgeb S, Moser M-B, Moser EI. 2007. Pattern separation in the dentate gyrus and CA3 of the hippocampus. *Science* 315: 961–6.
- Leutgeb S, Leutgeb JK. 2007. Pattern separation, pattern completion, and new neuronal codes within a continuous CA3 map. *Learn Mem* 14: 745–57.
- Leutgeb S, Leutgeb JK, Treves A, Moser M-B, Moser EI. 2004. Distinct ensemble codes in hippocampal areas CA3 and CA1. *Science* 305: 1295–8.
- Lim C, Mufson EJ, Kordower JH, Blume HW, Madsen JR, Saper CB. 1997. Connections of the hippocampal formation in humans: II. The endfolial fiber pathway. *J Comp Neurol* 385.
- Lindenberger U, Mayr U. 2014. Cognitive aging: is there a dark side to environmental support? *Trends Cogn Sci* 18: 7–15.
- Liu KY, Gould RL, Coulson MC, Ward E V., Howard RJ. 2016. Tests of pattern separation and pattern completion in humans-A systematic review. *Hippocampus* 26: 705–717.
- Loftus GR. 1972. Eye fixations and recognition memory for pictures. *Cogn Psychol* 3: 525–551.
- Lorente De Nó R. 1934. Studies on the structure of the cerebral cortex. II. Continuation of the study of the ammonic system. *J für Psychol und Neurol*.
- Luis CA, Keegan AP, Mullan M. 2009. Cross validation of the Montreal Cognitive Assessment in community dwelling older adults residing in the Southeastern US. *Int J Geriatr Psychiatry* 24: 197–201.
- Luo L, Craik FIM. 2008. Aging and memory: a cognitive approach. *Can J Psychiatry* 53: 346–53.
- Ly M, Murray EA, Yassa MA. 2013. Perceptual versus conceptual interference and pattern separation of verbal stimuli in young and older adults. *Hippocampus* 23: 425–30.
- Maass A, Berron D, Libby LA, Ranganath C, Düzel E. 2015. Functional subregions of the human entorhinal cortex. *Elife* 4: 1–20.
- Maass A, Schütze H, Speck O, Yonelinas AP, Tempelmann C, Heinze H-J, Berron D, Cardenas-Blanco A, Brodersen KH, Enno Stephan K, et al. 2014. Laminar activity in the hippocampus and entorhinal cortex related to novelty and episodic encoding. *Nat Commun* 5: 5547.
- Mai JK, Majtanik M, Paxinos G. 2015. *Atlas of the human brain*. Elsevier, Oxford.
- Malach R, Reppas JB, Benson RR, Kwong KK, Jiang H, Kennedy W a, Ledden PJ, Brady TJ, Rosen BR, Tootell RB. 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc Natl Acad Sci U S A* 92: 8135–8139.
- Malykhin N V., Lebel RM, Coupland NJ, Wilman AH, Carter R. 2010. In vivo quantification of hippocampal subfields using 4.7 T fast spin echo imaging. *Neuroimage* 49: 1224–1230.
- Marr D. 1971. Simple Memory: A Theory for Archicortex. *Philos Trans R Soc Lond B Biol Sci* 262: 23–81.

- McClelland JL. 1994. The organization of memory. A parallel distributed processing perspective. *Rev Neurol (Paris)* 150: 570–9.
- McClelland JL, Goddard NH. 1996. Considerations arising from a complementary learning systems perspective on hippocampus and neocortex. *Hippocampus* 6: 654–65.
- McClelland JL, McNaughton BL, O'Reilly RC, Barnes C, Becker S, Behrmann M, Myers CE, Plaut D, Skaggs W, Stark CEL, et al. 1995. Why There Are Complementary Learning Systems in the Hippocampus and Neocortex: Insights From the Successes and Failures of Connectionist Models of Learning and Memory. *Psychol Rev* 102: 419–457.
- McHugh TJ, Jones MW, Quinn JJ, Balthasar N, Coppari R, Elmquist JK, Lowell BB, Fanselow MS, Wilson MA, Tonegawa S. 2007. Dentate gyrus NMDA receptors mediate rapid pattern separation in the hippocampal network. *Science* 317: 94–9.
- McNaughton BL, Morris RGM. 1987. Hippocampal synaptic enhancement and information storage within a distributed memory system. *Trends Neurosci* 10: 408–415.
- Molitor RJ, Ko PC, Hussey EP, Ally B a. 2014. Memory-related eye movements challenge behavioral measures of pattern completion and pattern separation. *Hippocampus* 24: 666–72.
- Moschner C, Baloh RW. 1994. Age-related changes in visual tracking. *Journals Gerontol Med Sci* 49: M235-8.
- Motley SE, Kirwan CB. 2012. A parametric investigation of pattern separation processes in the medial temporal lobe. *J Neurosci* 32: 13076–85.
- Muehlhan M, Lueken U, Wittchen H-U, Kirschbaum C. 2011. The scanner as a stressor: Evidence from subjective and neuroendocrine stress parameters in the time course of a functional magnetic resonance imaging session. *Int J Psychophysiol* 79: 118–126.
- Mueller SG, Stables L, Du AT, Schuff N, Truran D, Cashdollar N, Weiner MW. 2007. Measurement of hippocampal subfields and age-related changes with high resolution MRI at 4T. *Neurobiol Aging* 28: 719–726.
- Mueller SG, Weiner MW. 2009. Selective effect of age, Apo e4, and Alzheimer's disease on hippocampal subfields. *Hippocampus* 19: 558–564.
- Mumford JA, Davis T, Poldrack RA. 2014. The impact of study design on pattern estimation for single-trial multivariate pattern analysis. *Neuroimage* 103: 130–138.
- Mumford JA, Turner BO, Ashby FG, Poldrack RA. 2012. Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *Neuroimage* 59: 2636–43.
- Myers CE, Scharfman HE. 2009. A role for hilar cells in pattern separation in the dentate gyrus: A computational approach. *Hippocampus* 19: 321–337.
- Myers CE, Scharfman HE. 2011. Pattern separation in the dentate gyrus: a role for the CA3 backprojection. *Hippocampus* 21: 1190–215.
- Nakashiba T, Cushman JD, Pelkey KA, Renaudineau S, Buhl DL, McHugh TJ, Barrera VR, Chittajallu R, Iwamoto KS, McBain CJ, et al. 2012. Young dentate granule cells mediate pattern separation, whereas old granule cells facilitate pattern completion. *Cell* 149: 188–201.
- Nakazawa K, Quirk MC, Chitwood RA, Watanabe M, Yeckel MF, Sun LD, Kato A, Carr CA, Johnston D, Wilson MA, et al. 2002. Requirement for hippocampal CA3 NMDA receptors in associative memory recall. *Science* 297: 211–8.
- Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, Chertkow H. 2005. The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *J Am Geriatr Soc* 53: 695–699.
- Neunuebel JP, Knierim JJ. 2014. CA3 Retrieves Coherent Representations from Degraded Input: Direct Evidence for CA3 Pattern Completion and Dentate Gyrus Pattern Separation. *Neuron* 81: 416–427.
- Nicolle MM, Gallagher M, McKinney M. 2001. Visualization of muscarinic receptor-mediated phosphoinositide turnover in the hippocampus of young and aged, learning-impaired long evans rats. *Hippocampus* 11: 741–746.
- O'Keefe J, Dostrovsky J. 1971. The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Res* 34: 171–175.
- O'Keefe J, Nadel L. 1978. *The hippocampus as a cognitive map*. Oxford University Press.
- O'Reilly RC, McClelland JL. 1994. Hippocampal conjunctive encoding, storage, and recall: avoiding a trade-off. *Hippocampus* 4: 661–82.
- O'Reilly RC, Norman KA, McClelland JL. 1998. A Hippocampal Model of Recognition Memory. In *Neural Information Processing Systems* (eds. M.I. Jordan, M.J. Kearns, and S.A. Solla), Vol. 10 of, pp. 73–79, MIT Press, Cambridge MA.
- O'Shea A, Cohen RA, Porges EC, Nissim NR, Woods AJ. 2016. Cognitive aging and the hippocampus in older adults. *Front Aging Neurosci* 8: 1–8.

- Olsen RK, Palombo DJ, Rabin JS, Levine B, Ryan JD, Rosenbaum RS. 2013. Volumetric analysis of medial temporal lobe subregions in developmental amnesia using high-resolution magnetic resonance imaging. *Hippocampus* 23: 855–860.
- Osorio-Gómez D, Guzmán-Ramos K, Bermúdez-Rattoni F. 2017. Memory trace reactivation and behavioral response during retrieval are differentially modulated by amygdalar glutamate receptors activity: interaction between amygdala and insular cortex. *Learn Mem* 24: 14–23.
- Paleja M, Girard T a., Christensen BK. 2011. Virtual human analogs to rodent spatial pattern separation and completion memory tasks. *Learn Motiv* 42: 237–244.
- Paleja M, Spaniol J. 2013. Spatial pattern completion deficits in older adults. *Front Aging Neurosci* 5: 3.
- Palombo DJ, Amaral RSC, Olsen RK, Muller DJ, Todd RM, Anderson AK, Levine B. 2013. KIBRA Polymorphism Is Associated with Individual Differences in Hippocampal Subregions: Evidence from Anatomical Segmentation using High-Resolution MRI. *J Neurosci* 33: 13088–13093.
- Parekh MB, Rutt BK, Purcell R, Chen Y, Zeineh MM. 2015. Ultra-high resolution in-vivo 7.0T structural imaging of the human hippocampus reveals the endfolial pathway. *Neuroimage* 112: 1–6.
- Park DC, Festini SB. 2016. Theories of Memory and Aging: A Look at the Past and a Glimpse of the Future. *J Gerontol B Psychol Sci Soc Sci* 72: gbw066.
- Park DC, Lautenschlager G, Hedden T, Davidson NS, Smith AD, Smith PK. 2002. Models of visuospatial and verbal memory across the adult life span. *Psychol Aging* 17: 299–320.
- Pertsov Y, Avidan G, Zohary E. 2009. Accumulation of visual information across multiple fixations. *J Vis* 9(10): 1–12.
- Poppenk JL, Evensmoen HR, Moscovitch M, Nadel L. 2013. Long-axis specialization of the human hippocampus. *Trends Cogn Sci* 17: 230–40.
- Prasad KMR, Patel AR, Muddasani S, Sweeney J, Keshavan MS. 2004. The Entorhinal Cortex in First-Episode Psychotic Disorders: A Structural Magnetic Resonance Imaging Study. *Am J Psychiatry* 161: 1612–1619.
- Preston AR, Bornstein AM, Hutchinson JB, Gaare ME, Glover GH, Wagner AD. 2010. High-resolution fMRI of Content-sensitive Subsequent Memory Responses in Human Medial Temporal Lobe. *J Cogn Neurosci* 22: 156–173.
- Pruessner JC, Köhler S, Crane J, Pruessner M, Lord C, Byrne A, Kabani N, Collins DL, Evans AC. 2002. Volumetry of temporopolar, perirhinal, entorhinal and parahippocampal cortex from high-resolution MR images: considering the variability of the collateral sulcus. *Cereb cortex* 12: 1342–1353.
- Ranganath C, Ritchey M. 2012. Two cortical systems for memory-guided behaviour. *Nat Rev Neurosci* 13: 713–26.
- Raz N, Lindenberger U, Rodrigue KM, Kennedy KM, Head D, Williamson A, Dahle C, Gerstorf D, Acker JD. 2005. Regional brain changes in aging healthy adults: General trends, individual differences and modifiers. *Cereb Cortex* 15: 1676–1689.
- Reagh ZM, Ho HD, Leal SL, Noche JA, Chun A, Murray EA, Yassa MA. 2015. Greater loss of object than spatial mnemonic discrimination in aged adults. *Hippocampus*.
- Reagh ZM, Roberts JM, Ly M, Diprospero N, Murray EA, Yassa MA. 2013. Spatial discrimination deficits as a function of mnemonic interference in aged adults with and without memory impairment. *Hippocampus* 0.
- Reagh ZM, Yassa MA. 2014a. Object and spatial mnemonic interference differentially engage lateral and medial entorhinal cortex in humans. *Proc Natl Acad Sci U S A* 111: E4264–E4273.
- Reagh ZM, Yassa MA. 2014b. Repetition strengthens target recognition but impairs similar lure discrimination: evidence for trace competition. *Learn Mem* 21: 342–6.
- Rey A. 1941. L'examen psychologique dans les cas d'encéphalopathie traumatique (Les problèmes). *Arch Psychol (Geneve)* 28: 215–285.
- Roberts JM, Ly M, Murray EA, Yassa MA. 2014. Temporal discrimination deficits as a function of lag interference in older adults. *Hippocampus* 8: 1–8.
- Rolls ET. 2016. Pattern separation, completion, and categorisation in the hippocampus and neocortex. *Neurobiol Learn Mem* 129: 4–28.
- Rolls ET. 2013. The mechanisms for pattern completion and pattern separation in the hippocampus. *Front Syst Neurosci* 7: 74.
- Rolls ET, Kesner RP. 2006. A computational theory of hippocampal function, and empirical tests of the theory. *Prog Neurobiol* 79: 1–48.
- Rolls ET, Kesner RP. 2016. Pattern separation and pattern completion in the hippocampal system. Introduction to the Special Issue. *Neurobiol Learn Mem* 129: 1–3.

- Rose M, Schmid C, Winzen A, Sommer T, Büchel C. 2005. The functional and temporal characteristics of top-down modulation in visual selection. *Cereb Cortex* 15: 1290–1298.
- Rosenbaum RS, Gilboa A, Moscovitch M. 2014. Case studies continue to illuminate the cognitive neuroscience of memory. *Ann N Y Acad Sci* 1316: 105–133.
- Rusinek H, Endo Y, Santi DS, Frid D, Tsui -H W, Segal S, Convit A, de Leon MJ. 2004. Atrophy rate in medial temporal lobe during progression of Alzheimer disease. 63: 2354–2359.
- Sahay A, Hen R. 2007. Adult hippocampal neurogenesis in depression. *Nat Neurosci* 10: 1110–1115.
- Schacter DL, Koutstaal W, Norman KA. 1997. False memories and aging. *Trends Cogn Sci* 1: 229–236.
- Scharfman HE. 2007. The CA3 “backprojection” to the dentate gyrus. *Prog Brain Res* 163: 627–637.
- Schlichting ML, Zeithamova D, Preston AR. 2014. CA1 subfield contributions to memory integration and inference. *Hippocampus* 24: 1248–1260.
- Sengupta A, Yakupov R, Speck O, Pollmann S, Hanke M. 2017. The effect of acquisition resolution on orientation decoding from V1 BOLD fMRI at 7T. *Neuroimage* 148: 64–76.
- Shing YL, Rodrigue KM, Kennedy KM, Fandakova Y, Bodammer NC, Werkle-Bergner M, Lindenberger U, Raz N. 2011. Hippocampal subfield volumes: age, vascular risk, and correlation with associative memory. *Front Aging Neurosci* 3: 2.
- Shrout P, Fleiss J. 1979. Intraclass correlations: uses in assessing rater reliability. 86: 420–8.
- Simić G, Kostović I, Winblad B, Bogdanović N. 1997. Volume and number of neurons of the human hippocampal formation in normal aging and Alzheimer’s disease. *J Comp Neurol* 379: 482–94.
- Slotnick SD. 2013. The nature of recollection in behavior and the brain. *Neuroreport* 24: 663–70.
- Small SA. 2014. Isolating Pathogenic Mechanisms Embedded within the Hippocampal Circuit through Regional Vulnerability. *Neuron* 84: 32–39.
- Small SA, Chawla MK, Buonocore M, Rapp PR, Barnes CA. 2004. Imaging correlates of brain function in monkeys and rats isolates a hippocampal subregion differentially vulnerable to aging. *Proc Natl Acad Sci U S A* 101: 7181–6.
- Small SA, Schobel S a, Buxton RB, Witter MP, Barnes C a. 2011. A pathophysiological framework of hippocampal dysfunction in ageing and disease. *Nat Rev Neurosci* 12: 585–601.
- Smith CN, Wixted JT, Squire LR. 2011. The Hippocampus Supports Both Recollection and Familiarity When Memories Are Strong. *J Neurosci* 31: 15693–15702.
- Smith TD, Adams MM, Gallagher M, Morrison JH, Rapp PR. 2000. Circuit-specific alterations in hippocampal synaptophysin immunoreactivity predict spatial learning impairment in aged rats. *J Neurosci* 20: 6587–93.
- Squire LR. 2004. Memory systems of the brain: A brief history and current perspective. *Neurobiol Learn Mem* 82: 171–177.
- Squire LR, Cohen NJ, Nadel L. 1984. The medial temporal region and memory consolidation: A new hypothesis. *Mem Consol Psychobiol Cogn* 185–210.
- Squire LR, Stark CEL, Clark RE. 2004. The medial temporal lobe. *Annu Rev Neurosci* 27: 279–306.
- Squire LR, Zola-Morgan S. 1991. The Medial Temporal Lobe Memory System. *Science (80- )* 253: 1380–1386.
- Stanley DP, Shetty AK. 2004. Aging in the rat hippocampus is associated with widespread reductions in the number of glutamate decarboxylase-67 positive interneurons but not interneuron degeneration. *J Neurochem* 89: 204–216.
- Staresina BP, Alink A, Kriegeskorte N, Henson RN. 2013a. Awake reactivation predicts memory in humans. *Proc Natl Acad Sci U S A* 110: 21159–21164.
- Staresina BP, Cooper E, Henson RN. 2013b. Reversible Information Flow across the Medial Temporal Lobe: The Hippocampus Links Cortical Modules during Memory Retrieval. *J Neurosci* 33: 14184–14192.
- Staresina BP, Henson RN, Kriegeskorte N, Alink A. 2012. Episodic reinstatement in the medial temporal lobe. *J Neurosci* 32: 18150–6.
- Stark SM, Stevenson R, Wu C, Rutledge S, Stark CEL. 2015. Stability of age-related deficits in the mnemonic similarity task across task variations. *Behav Neurosci* 129: 257–268.
- Stark SM, Yassa MA, Lacy JW, Stark CEL. 2013. A task to assess behavioral pattern separation (BPS) in humans: Data from healthy aging and mild cognitive impairment. *Neuropsychologia* 51: 2442–9.
- Steward O. 1976. Topographic organization of the projections from the entorhinal area to the hippocampal formation of the rat. *J Comp Neurol* 167: 285–314.
- Stöcker T, Shah NJ. 2006. MP-SAGE: A new MP-RAGE sequence with enhanced SNR and CNR for brain imaging utilizing square-spiral phase encoding and variable flip angles. *Magn Reson Med* 56: 824–834.

- Stokes JD, Kyle CT, Ekstrom AD. 2015. Complementary Roles of Human Hippocampal Subfields in Differentiation and Integration of Spatial Context. *J Cogn Neurosci* 27: 546–59.
- Suthana NA, Donix M, Wozny DR, Bazih AJ, Jones M, Heidemann RM, Trampel R, Ekstrom AD, Scharf M, Knowlton BJ, et al. 2015. High-resolution 7T fMRI of Human Hippocampal Subfields during Associative Learning. *J Cogn Neurosci* 27: 1194–1206.
- Toner CK, Pirogovsky E, Kirwan CB, Gilbert PE. 2009. Visual object pattern separation deficits in nondemented older adults. *Learn Mem* 16: 338–42.
- Travis SG, Huang Y, Fujiwara E, Radomski A, Olsen F, Carter R, Seres P, Malykhin N V. 2014. High field structural MRI reveals specific episodic memory correlates in the subfields of the hippocampus. *Neuropsychologia* 53: 233–245.
- Treves A, Rolls ET. 1994. Computational analysis of the role of the hippocampus in memory. *Hippocampus* 4: 374–391.
- van Veluw SJ, Wisse LEM, Kuijf HJ, Spliet WGM, Hendrikse J, Luijten PR, Geerlings MI, Biessels GJ. 2013. Hippocampal T2 hyperintensities on 7Tesla MRI. *NeuroImage Clin* 3: 196–201.
- Vann SD, Aggleton JP, Maguire EA. 2009. What does the retrosplenial cortex do? *Nat Rev Neurosci* 10: 792–802.
- Vazdarjanova A, Guzowski JF. 2004. Differences in hippocampal neuronal population responses to modifications of an environmental context: evidence for distinct, yet complementary, functions of CA3 and CA1 ensembles. *J Neurosci* 24: 6489–96.
- Vela J, Gutierrez A, Vitorica J, Ruano D. 2003. Rat hippocampal GABAergic molecular markers are differentially affected by ageing. *J Neurochem* 85: 368–377.
- Walther A, Nili H, Ejaz N, Alink A, Kriegeskorte N, Diedrichsen J. 2015. Reliability of dissimilarity measures for multi-voxel pattern analysis. *Neuroimage*.
- Wang Z, Neylan TC, Mueller SG, Lenoci M, Truran D, Marmar CR, Weiner MW, Schuff N. 2010. Magnetic Resonance Imaging of Hippocampal Subfields in Posttraumatic Stress Disorder. *Arch Gen Psychiatry* 67: 296.
- Wechsler D. 2008. *Wechsler Adult Intelligence Scale-Fourth Edition*. The Psychological Corporation, San Antonio, TX.
- Wilson IA, Gallagher M, Eichenbaum H, Tanila H. 2006. Neurocognitive aging: prior memories hinder new hippocampal encoding. *Trends Neurosci* 29: 662–70.
- Wilson IA, Ikonen S, Gallagher M, Eichenbaum H, Tanila H. 2005. Age-associated alterations of hippocampal place cells are subregion specific. *J Neurosci* 25: 6877–86.
- Winterburn JL, Pruessner JC, Chavez S, Schira MM, Lobaugh NJ, Voineskos AN, Chakravarty MM. 2013. A novel in vivo atlas of human hippocampal subfields using high-resolution 3 T magnetic resonance imaging. *Neuroimage* 74: 254–65.
- Wisse LEM, Adler DH, Ittyerah R, Pluta JB, Robinson JL, Schuck T, Trojanowski JQ, Grossman M, Detre JA, Elliott MA, et al. 2016a. Comparison of In Vivo and Ex Vivo MRI of the Human Hippocampal Formation in the Same Subjects. *Cereb Cortex* Epub: 1–12.
- Wisse LEM, Daugherty AM, Olsen RK, Berron D, Carr VA, Stark CEL, Amaral RSC, Amunts K, Augustinack JC, Bender AR, et al. 2016b. A harmonized segmentation protocol for hippocampal and parahippocampal subregions: Why do we need one and what are the key goals? *Hippocampus* 27: 3–11.
- Wisse LEM, Gerritsen L, Zwanenburg JJM, Kuijf HJ, Luijten PR, Biessels GJ, Geerlings MI. 2012. Subfields of the hippocampal formation at 7 T MRI: in vivo volumetric assessment. *Neuroimage* 61: 1043–9.
- Wisse LEM, Kuijf HJ, Honingh AM, Wang H, Pluta JB, Das SR, Wolk DA, Zwanenburg JJM, Yushkevich PA, Geerlings MI. 2016c. Automated Hippocampal Subfield Segmentation at 7T MRI. *Am J Neuroradiol* 37: 1050–1057.
- Wixted JT, Squire LR. 2010. The role of the human hippocampus in familiarity-based and recollection-based recognition memory. *Behav Brain Res* 215: 197–208.
- Wolbers T, Büchel C. 2005. Dissociable Retrosplenial and Hippocampal Contributions to Successful Formation of Survey Representations. *J Neurosci* 25: 3333–3340.
- Wolk DA, Das SR, Mueller SG, Weiner MW, Yushkevich PA. 2017. Medial temporal lobe subregional morphometry using high resolution MRI in Alzheimer's disease. *Neurobiol Aging* 49: 204–213.
- Xie L, Pluta JB, Das SR, Wisse LEM, Wang H, Mancuso L, Kliot D, Avants BB, Ding S-L, Manjón J V., et al. 2017. Multi-template analysis of human perirhinal cortex in brain MRI: Explicitly accounting for anatomical variability. *Neuroimage* 144: 183–202.
- Xie L, Wisse LEM, Das SR, Wang H, Wolk DA, Manjón J V, Yushkevich PA. 2016. *Accounting for the Confound of Meninges in Segmenting Entorhinal and Perirhinal Cortices in T1-Weighted MRI*.
- Xu W, Südhof TC. 2013. A Neural Circuit for Memory Specificity and Generalization. *Science* 339.

- Yassa MA, Lacy JW, Stark SM, Albert MS, Gallagher M, Stark CEL. 2011a. Pattern separation deficits associated with increased hippocampal CA3 and dentate gyrus activity in nondemented older adults. *Hippocampus* 21: 968–79.
- Yassa MA, Mattfeld AT, Stark SM, Stark CEL. 2011b. Age-related memory deficits linked to circuit-specific disruptions in the hippocampus. *Proc Natl Acad Sci U S A* 108: 8873–8.
- Yassa MA, Stark CEL. 2011. Pattern separation in the hippocampus. *Trends Neurosci* 34: 515–525.
- Yonelinas AP. 2002. The Nature of Recollection and Familiarity: A Review of 30 Years of Research. *J Mem Lang* 46: 441–517.
- Yonelinas AP, Aly M, Wang WC, Koen JD. 2010. Recollection and familiarity: Examining controversial assumptions and new directions. *Hippocampus* 20: 1178–1194.
- Yushkevich PA, Amaral RSC, Augustinack JC, Bender AR, Bernstein JD, Boccardi M, Bocchetta M, Burggren AC, Carr V a., Chakravarty MM, et al. 2015a. Quantitative comparison of 21 protocols for labeling hippocampal subfields and parahippocampal subregions in in vivo MRI: Towards a harmonized segmentation protocol. *Neuroimage* 111: 526–541.
- Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, Gerig G. 2006. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage* 31: 1116–1128.
- Yushkevich PA, Pluta JB, Wang H, Xie L, Ding S-LL, Gertje EC, Mancuso L, Klot D, Das SR, Wolk DA. 2015b. Automated volumetry and regional thickness analysis of hippocampal subfields and medial temporal cortical structures in mild cognitive impairment. *Hum Brain Mapp* 36: 258–287.
- Yushkevich PA, Wang H, Pluta JB, Das SR, Craige C, Avants BB, Weiner MW, Mueller SG. 2010. Nearly automatic segmentation of hippocampal subfields in in vivo focal T2-weighted MRI. *Neuroimage* 53: 1208–1224.
- Zeidman P, Maguire EA. 2016. Anterior hippocampus: the anatomy of perception, imagination and episodic memory. *Nat Rev Neurosci* 17: 173–182.
- Zeineh MM, Engel SA, Thompson PM, Bookheimer SY. 2001. Unfolding the human hippocampus with high resolution structural and functional MRI. 265: 111–120.
- Zheng J, Anderson KL, Leal SL, Shestyuk A, Gulsen G, Mnatsakanyan L, Vadera S, Hsu FPK, Yassa MA, Knight RT, et al. 2017. Amygdala-hippocampal dynamics during salient information processing. *Nat Commun* 8: 14413.

# APPENDIX

## LIST OF ABBREVIATIONS

<b>ACh</b>	acetylcholine	<b>LOC</b>	lateral occipital complex
<b>AD</b>	Alzheimer's disease	<b>MCI</b>	mild cognitive impairment
<b>AG</b>	ambient gyrus	<b>MIC</b>	Memory Image Completion
<b>ANOVA</b>	analysis of variance	<b>MoCA</b>	Montreal Cognitive Assessment
<b>ANTs</b>	Advanced Normalization Tools	<b>MPS</b>	multi-voxel pattern similarity
<b>ASHS</b>	Automatic Segmentation of Hippocampal Subfields	<b>MPSAGE</b>	magnetization-prepared spiral acquisition gradient-echo
<b>BOLD</b>	blood-oxygen-level dependent	<b>MRI</b>	magnetic resonance imaging
<b>CA1</b>	Cornu Ammonis 1	<b>MS</b>	medial septum
<b>CA2</b>	Cornu Ammonis 2	<b>MST</b>	Mnemonic Similarity Task
<b>CA3</b>	Cornu Ammonis 3	<b>MTL</b>	medial temporal lobe
<b>CA4</b>	Cornu Ammonis 4	<b>MVPA</b>	multivariate pattern analysis
<b>CaS</b>	calcarine sulcus	<b>MWT-B</b>	multiple-choice word test
<b>CR</b>	correct rejection	<b>NMDA</b>	N-Methyl-D-aspartate
<b>CS</b>	collateral sulcus	<b>OTS</b>	occipito-temporal sulcus
<b>CSa</b>	anterior	<b>PhC</b>	parahippocampal cortex
<b>CSp</b>	posterior	<b>PhG</b>	parahippocampal gyrus
<b>CSF</b>	cerebrospinal fluid	<b>PN</b>	partial new
<b>DA</b>	dopamine	<b>PNc</b>	partial new correct
<b>DG</b>	dentate gyrus	<b>PNf</b>	partial new false
<b>DR</b>	delayed recall	<b>PO</b>	partial old
<b>DSI</b>	Dice Similarity Index	<b>POc</b>	partial old correct
<b>DSST</b>	Digit Symbol Substitution Test	<b>POfn</b>	partial old false identified as new
<b>EPI</b>	echo-planar images	<b>POfo</b>	partial old false identified as other old
<b>ErC</b>	entorhinal cortex	<b>PrC</b>	perirhinal cortex
<b>FA</b>	false alarm	<b>ROCF</b>	Rey-Osterrieth Complex Figure
<b>FDR</b>	false discovery rate	<b>ROI</b>	region-of-interest
<b>FG</b>	fusiform gyrus	<b>SaS</b>	semiannular sulcus
<b>fMRI</b>	functional magnetic resonance imaging	<b>SE</b>	standard error
<b>FO</b>	full old	<b>SPM</b>	Statistical Parametric Mapping
<b>FOV</b>	field-of-view	<b>SRLM</b>	stratum radiatum lacunosum moleculare
<b>FSL</b>	FMRIB Software Library	<b>STS</b>	superior temporal sulcus
<b>FWE</b>	family-wise error	<b>Sub</b>	subiculum
<b>GLM</b>	general linear model	<b>T</b>	Tesla
<b>HB</b>	hippocampal body	<b>TE</b>	echo time
<b>HH</b>	hippocampal head	<b>TI</b>	inversion time
<b>HRF</b>	hemodynamic response function	<b>TR</b>	repetition time
<b>HT</b>	hippocampal tail	<b>TSE</b>	turbo spin echo
<b>ICC</b>	intra-class correlation coefficient	<b>VTA</b>	ventral tegmental area
<b>int</b>	interneuron		
<b>ISI</b>	interstimulus interval		
<b>LDI</b>	Lure Discrimination Index		



## LIST OF FIGURES

Figure 1. Hippocampal connections.....	2
Figure 2. Hippocampal degeneration in aging.....	9
Figure 3. Ch. 2 - design of the test phase.....	15
Figure 4. Ch. 2 - performance and bias measures.....	17
Figure 5. Ch. 2 - response distribution for new stimuli.....	19
Figure 6. Ch. 2 - reaction times.....	19
Figure 7. Ch. 2 - confidence ratings.....	20
Figure 8. Ch. 3 - performance and bias measures.....	29
Figure 9. Ch. 3 - response distribution for new stimuli.....	29
Figure 10. Ch. 3 - reaction times.....	30
Figure 11. Ch. 3 - confidence ratings.....	31
Figure 12. Ch. 3 - eye-tracking data for the whole images at study and test.....	32
Figure 13. Ch. 3 - eye-tracking data for the regions-of-interest (ROIs).....	33
Figure 14. Ch. 4 - Performance and bias measures.....	37
Figure 15. Ch. 4 - Performance on the Mnemonic similarity task (MST).....	38
Figure 16. Ch. 4 - Bias scores for the Mnemonic similarity task (MST) and the Memory Image Completion task (MIC).....	38
Figure 17. Ch. 5 - experimental design of the test phase.....	43
Figure 18. Ch. 5 – MRI volumes and segmentations.....	45
Figure 19. Ch. 5 – Single trial correlation matrices dependent on model specifications.....	46
Figure 20. Ch. 5 – Response-based analysis design.....	47
Figure 21. Ch. 5 – Performance measures.....	51
Figure 22. Ch. 5 - Performance measures by session.....	52
Figure 23. Ch. 5 – Proof of concept correlations.....	53
Figure 24. Ch. 5 – Pattern completion effects.....	54
Figure 25. Ch. 5 – Retrieval effects in young adults only.....	54
Figure 26. Ch. 5 – Novelty effects in young adults only.....	55
Figure 27. Ch. 5 – Visibility effects in both age groups.....	55
Figure 28. Ch. 5 – Age effects across the whole task.....	56
Figure 29: Segmentation hierarchy.....	68
Figure 30: Excluded structures in a coronal view.....	69
Figure 31: Slice-by-slice segmentation for a type 1 collateral sulcus (CS) – anterior part.....	71
Figure 32: Continuation of Figure 31 - Slice-by-slice segmentation for a type 1 collateral sulcus (CS) – posterior part.....	72
Figure 33: Different depths of the collateral sulcus (CS) with respective segmentation rules applied.....	75
Figure 34: Slice-by-slice segmentation for a type II collateral sulcus (CS) – anterior part.....	77
Figure 35: Rules for hippocampal subfield segmentation shown on the relevant slices from anterior to posterior.....	80
Figure 36: Heuristic rules for separation of DG and CA3 if the endfolial pathway is not visible.....	82
Figure 37: Exemplary segmentation profile.....	83

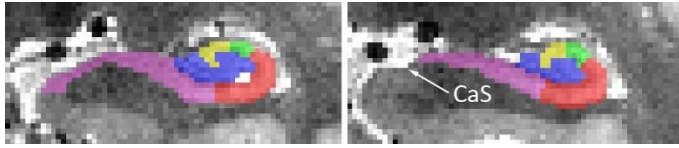
## LIST OF TABLES

Table 1. Ch. 2 - false alarm rates for learned stimuli. ....	17
Table 2. Ch. 3 - health questionnaire and neuropsychological data. ....	26
Table 3. Ch. 3 - false alarm rates for learned stimuli. ....	30
Table 4. Ch. 5 - stimulus material for the test phase. ....	42
Table 5. Ch. 5 - false alarm rates for learned stimuli. ....	51
Table 6. Ch. 5 – Brain regions showing significant effects in young adults. ....	56
Table 7. Ch. 5 – Brain regions showing significant effects in older adults. ....	57
Table 8. Ch. 5 – Brain regions showing significant effects between age groups across the whole task. ...	58
Table 9. Intra-rater reliability for all subregions. ....	85
Table 10. Inter-rater reliability in type 1 and 2 CS patterns separately. ....	86
Table 11. Inter-rater reliability for all subregions. ....	87
Table 12. Volumes of all regions in ml in comparison to other studies. ....	88

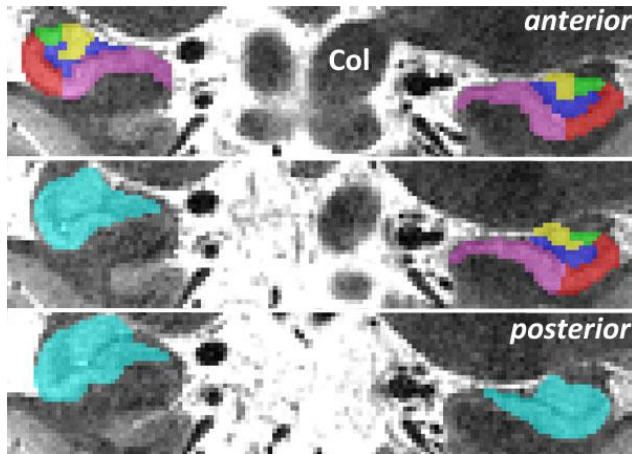
## LIST OF SUPPLEMENTARY FIGURES

Supplementary Figure 1: Calcarine sulcus (CaS). ....	V
Supplementary Figure 2: Colliculi. ....	V
Supplementary Figure 3: Sulcus depth measurement. ....	V
Supplementary Figure 4: Decision tree for segmentation of area 35. ....	V
Supplementary Figure 5: Double occipito-temporal sulcus (OTS). ....	VI
Supplementary Figure 6: Mid-fusiform sulcus. ....	VI
Supplementary Figure 7: Rhinal sulcus. ....	VI
Supplementary Figure 8: Parahippocampal-ligual sulcus (PhligS) and calcarine sulcus (CaS). ....	VII
Supplementary Figure 9: Gap between anterior and posterior collateral sulcus (CS). ....	VII
Supplementary Figure 10: Post segmentation checklist. ....	VIII
Supplementary Figure 11: Bland Altman plots of area 35 volume for different sulcal patterns. ....	VIII

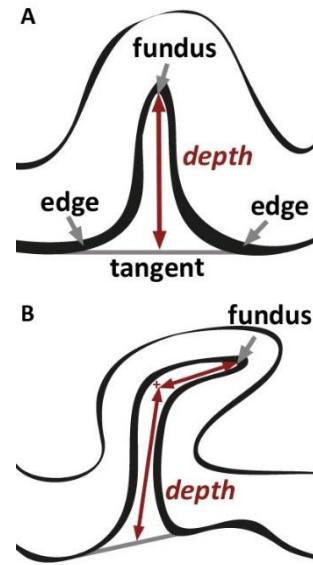
## SUPPLEMENTARY FIGURES



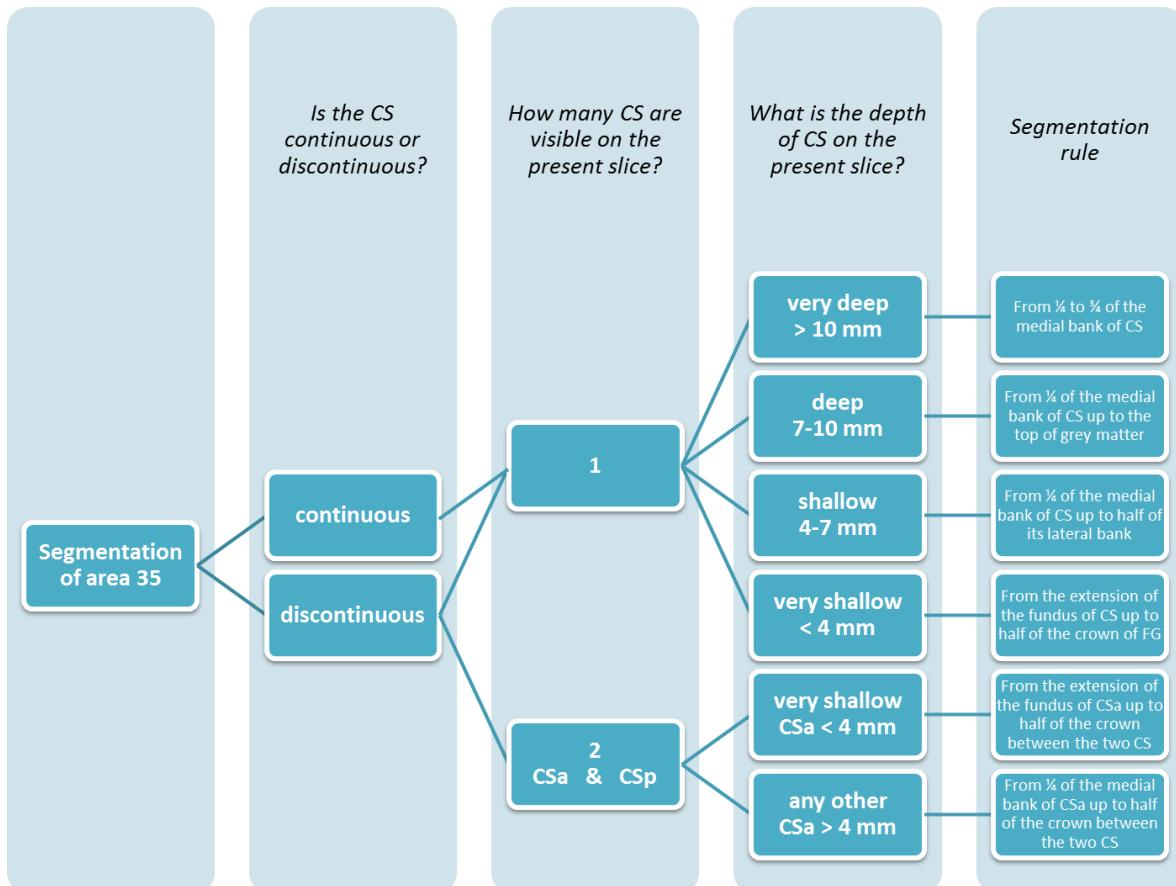
**Supplementary Figure 1: Calcarine sulcus (CaS).** When it appears, its banks are spared from segmentation.



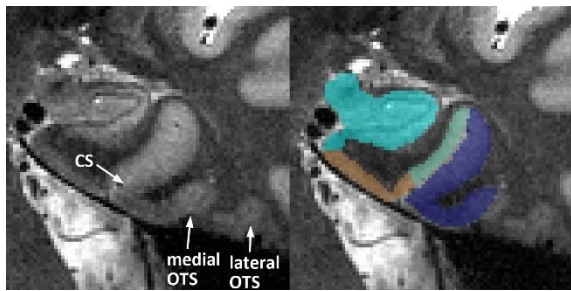
**Supplementary Figure 2: Colliculi.** Transition from hippocampal body (top) to hippocampal tail (bottom) dependent on the presence of the inferior and superior colliculi (Col) for each hemisphere separately. Hippocampal subfields are only segmented in the body, and not in the tail.



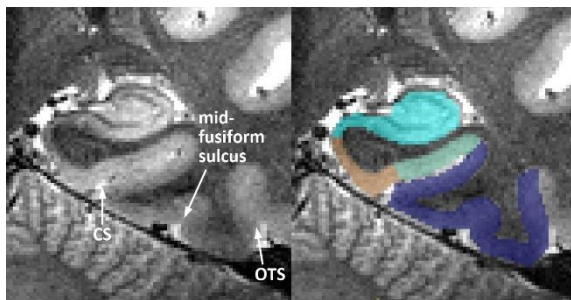
**Supplementary Figure 3: Sulcus depth measurement.** The edges adjacent to the sulcus are connected via a tangent line (grey line). (A) The depth of a straight sulcus is measured from the middle of the tangent to the fundus of the sulcus (red arrow). (B) If the sulcus bends, the depth is measured along the middle of the sulcus in separate legs, the lengths of which are summed up (red arrows connected via +).



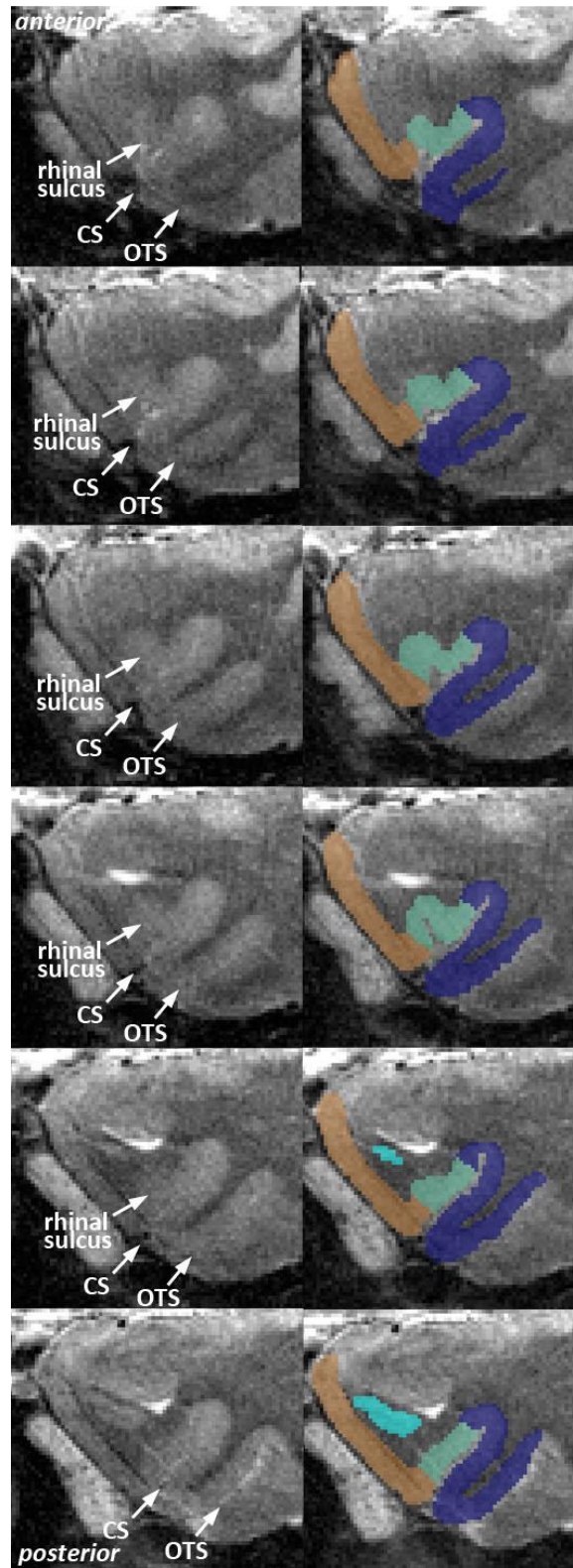
**Supplementary Figure 4: Decision tree for segmentation of area 35.** Applies to every coronal slice.



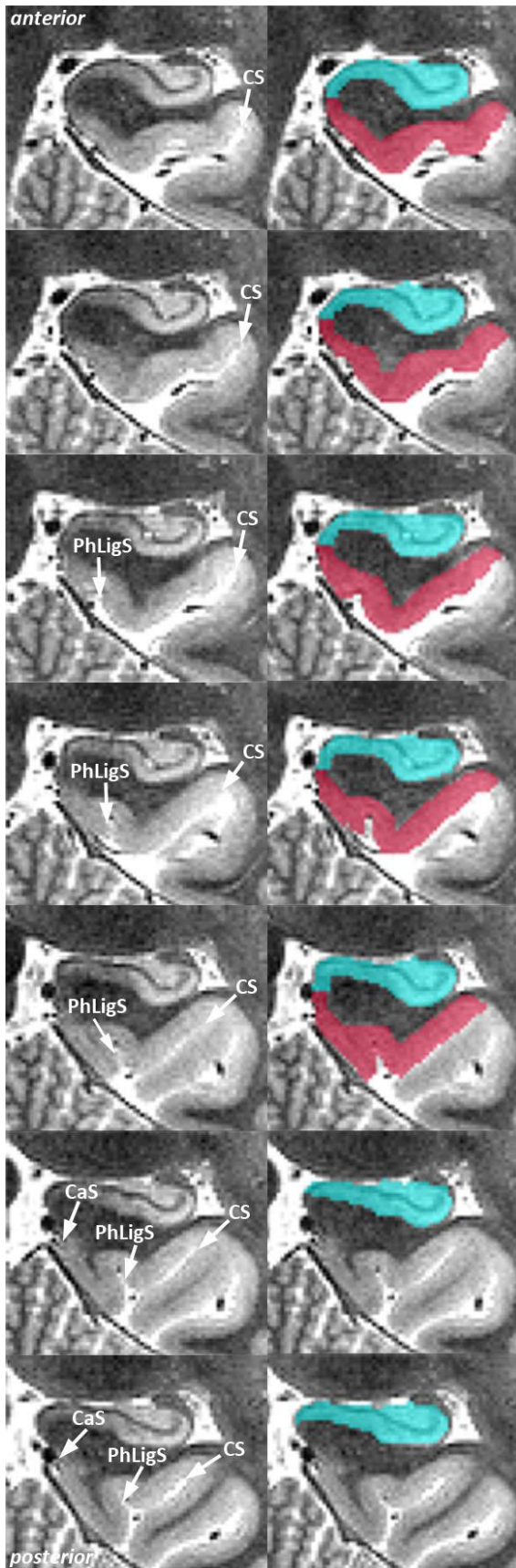
**Supplementary Figure 5: Double occipito-temporal sulcus (OTS).** This is an example of the variability of OTS. In case of two OTS or a bifurcated OTS, the more medial OTS is chosen as the lateral border of area 36.



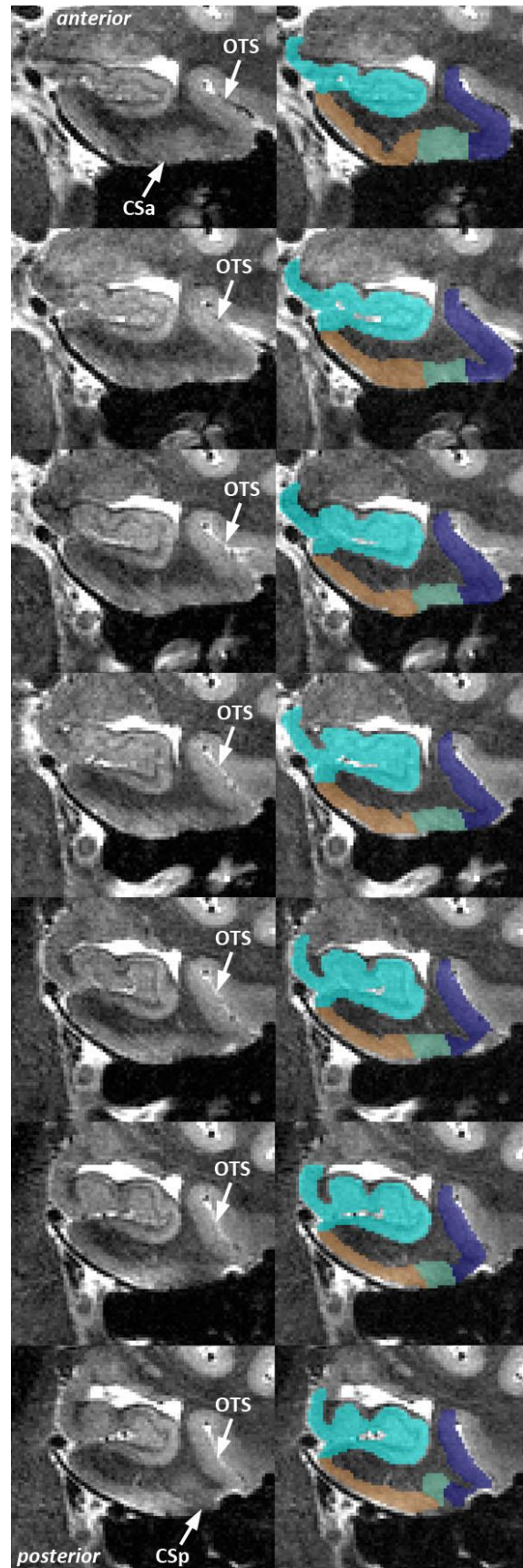
**Supplementary Figure 6: Mid-fusiform sulcus.** This sulcus lies between the collateral sulcus (CS) and the occipitotemporal sulcus (OTS). It can be identified as it is considerably shallower than OTS, and it usually appears only in very posterior slices; mostly it even appears after perirhinal cortex segmentation has stopped.



**Supplementary Figure 7: Rhinal sulcus.** It lies medial to the collateral sulcus (CS) and is more shallow; it often 'moves' up the CS. Usually, it is visible in very anterior slices, mostly even before segmentation starts. If the rhinal sulcus is separate from CS within the segmentation range the rules change (see text in 6.2.5.2), but this occurs only in very rare cases.



**Supplementary Figure 8: Parahippocampal-ligual sulcus (PhLigS) and calcarine sulcus (CaS).** The PhLigS often 'moves' laterally up along the collateral sulcus (CS). CaS is always medial to PhLigS.



**Supplementary Figure 9: Gap between anterior and posterior collateral sulcus (CS).** When neither CSa nor CSp are visible in the coronal slice, the boundaries between entorhinal cortex, area 35 and 36 should be extrapolated from the next slices were the collateral sulci can be identified.

## Post segmentation checklist

### For all structures:

- cysts, CSF, wide sulci, blood vessels, fimbria, alveus and meninges are excluded from segmentation

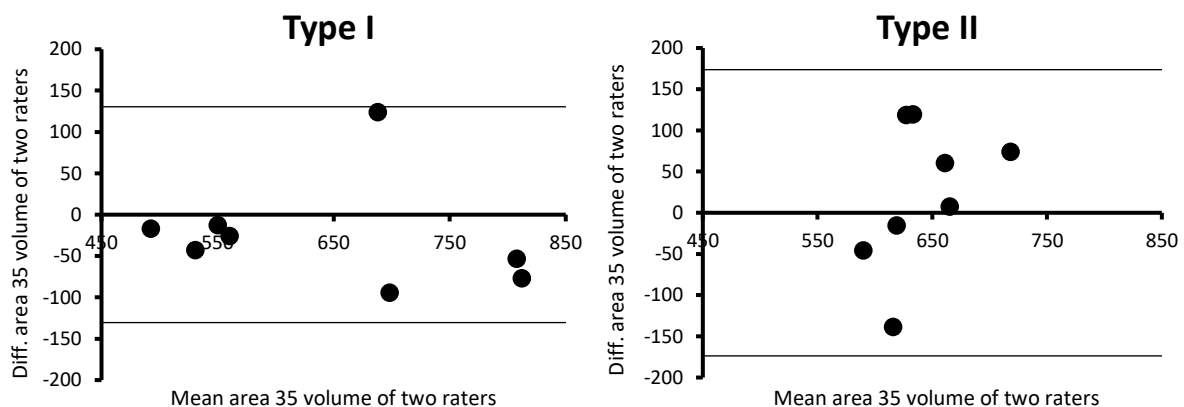
### Hippocampus:

- SRLM is equally divided between structures
- all borders are applied orthogonally to the structure
- HH → until uncus apex
  - (1) Sub and ErC are connected when the uncus sulcus can be followed from surface to its fundus
  - (2) DG replaces the SRLM as the most lateral point for CA1/Sub border when DG appears
  - (3) CA1 replaces Sub superiorly 2 slices (2.2 mm) posterior to where DG appears
  - (4) last 4 HH slices (4.4 mm) include CA2/CA3 segmentation
    - borders extrapolated from slice where the uncus is separate from the hippocampus
    - CA1/Sub border at  $\frac{1}{4}$  medial to lateral DG
- HB → when uncus has disappeared
  - (1) CA1/Sub border at  $\frac{1}{2}$  medial to lateral DG
  - (2) CaS is excluded from segmentation
  - (3) subfield segmentation stops when colliculi of the particular hemisphere disappear entirely

### Extrahippocampal regions:

- ErC/area 35/area 36 start 4 slices (4.4 mm) anterior to HH
- ErC superiorly ends at SaS (possibly extrapolated from posterior slices)
- Sulcus depths: see decision tree in Supplementary Figure 4
- ErC stops 2 slices (2.2 mm) posterior to HH (uncus apex)
- Area 35/36 stop 4 slices (4.4 mm) posterior to HH (uncus apex)
- Transitions:
  - (1) last slice of ErC
  - (2) on CSa slice for type II when CSa only changes to CSp only
  - (3) when OTS jumps
- PhC begins 5 slices (5.5 mm) posterior to HH directly after the end of area 35/area 36
- PhC stops when CaS appears

Supplementary Figure 10: Post segmentation checklist.



Supplementary Figure 11: Bland Altman plots of area 35 volume for different sulcal patterns.

# DECLARATION/ERKLÄRUNG

I herewith declare that this thesis entitled *Recognition memory associated with hippocampal pattern completion in young and older adults* is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references. Additionally, this work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Hiermit erkläre ich, dass ich die von mir eingereichte Dissertation zum Thema *Recognition memory associated with hippocampal pattern completion in young and older adults* selbständig verfasst, nicht bereits als Dissertation verwendet habe und die benutzten Hilfsmittel und Quellen vollständig angegeben wurden. Weiterhin erkläre ich, dass ich weder diese noch eine andere Arbeit zur Erlangung des akademischen Grades doctor rerum naturalium (Dr. rer. nat.) an anderen Einrichtungen eingereicht habe.

Magdeburg, 21.02.2017

Paula Vieweg

# PAULA VIEWEG | CURRICULUM VITAE

date of birth: 16<sup>th</sup> November 1987 | place of birth: Bautzen, Germany | nationality: German

## EDUCATION

- 04/2012 **German Center for Neurodegenerative Diseases (DZNE)**  
– 09/2017 **Otto-von-Guericke-University, Magdeburg, Germany**  
PhD in Neuroscience (summa cum laude)  
Supervisor: Prof. Dr. Thomas Wolbers  
*Recognition memory associated with hippocampal pattern completion in young and older adults*
- 08/2009 **University of Amsterdam, Netherlands**  
– 09/2011 MSc in Brain and Cognitive Science (with distinction)  
02/2011 Literature thesis, Supervisor: Jan Mehnert  
– 06/2011 *Near-infrared spectroscopy in motor imagery and movement therapy with brain-computer interfaces*
- 01/2011 **University of Edinburgh, UK, Centre for Cognitive & Neural Systems**  
– 07/2011 Research internship, Supervisor: Dr. Thomas Wolbers  
*Pattern separation and completion in humans*
- 02/2010 **VU Amsterdam, Netherlands, Department of Cognitive Psychology**  
– 10/2010 Research internship, Supervisor: Dr. Hannie van Hooff  
*Attentional differentiation of two negative emotions: fear and disgust*
- 10/2006 **University of Osnabrück, Germany**  
– 07/2009 BSc in Cognitive Science
- 01/2009 **University of Osnabrück, Germany, Institute of Cognitive Science**  
– 07/2009 Student assistant for EEG and eye-tracking in the Neurobiopsychology group (Prof. Dr. Peter König)
- 07/2008 **University of Otago, New Zealand, Department of Psychology**  
– 11/2008 Research Internship, Supervisor: Prof. Jeff Miller  
*Subjective reports of stimulus, response, and decision times in speeded tasks: How accurate are decision time reports?*
- 08/1998 **Friedrich-Schiller-Gymnasium Bautzen, Germany**  
– 07/2006 University-entrance diploma (Abitur)

## PUBLICATIONS

- Berron, D.\*, **Vieweg, P.\***, Hochkeppeler, A., Pluta, J.B., Ding, S.-L., Maass, A., Luther, A., Xie, L., Das, S.R., Wolk, D.A., Wolbers, T., Yushkevich, P.A.\*, Düzel, E.\*, Wisse, L.E.M.\* (2017). A protocol for manual segmentation of medial temporal lobe subregions in 7 tesla MRI. *NeuroImage: Clinical*, 15C, 466–482. [\*Denotes equal first and senior author contributions.]
- Baker, S., **Vieweg, P.**, Gao, F., Gilboa, A., Wolbers, T., Black, S.E., & Rosenbaum, R.S. (2016). The Human Dentate Gyrus Plays a Necessary Role in Discriminating New Memories. *Current Biology*, 26, 2629-2634.
- Vieweg, P.**, Stangl, M., Howard, L.R. & Wolbers, T. (2015). Changes in pattern completion - a key mechanism to explain age-related recognition memory deficits? *Cortex*, 64, 343-351.
- Van Hooff, J. C., Devue, C., **Vieweg, P. E.**, & Theeuwes, J. (2013). Disgust- and not fear-evoking images hold our attention. *Acta Psychologica*, 143(1), 1–6.
- Miller, J., **Vieweg, P.**, Kruize, N., & McLea, B. (2010). Subjective reports of stimulus, response, and decision times in speeded tasks: How accurate are decision time reports? *Consciousness and Cognition*, 19(4), 1013-1036.