Noticing the new: novelty encoding and ageing in visual immediate memory

Amy Victoria Smith

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

The University of Leeds

School of Psychology

June, 2019

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

The work in Chapter 2 (Experiment 1 and 2) has been included in a jointly authored publication as follows:

Smith, A. V., McKeown, D., & Bunce, D. (2017). Time manages interference in visual short term memory, *Memory*, *25*(8), 1117-1128. doi: 10.1080/09658211.2016.1270967

Used by permission from Memory, ©2017 Taylor & Francis Group
http://www.tandfonline.com doi: 10.1080/09658211.2016.1270967

Amy Smith designed the experiments and conducted all data collection and analysis for this chapter with the supervision of Dr Denis McKeown. The manuscripts were written by Amy Smith with guidance on drafts from Dr Denis McKeown and Professor David Bunce.

Acknowledgements

First and foremost I would like to thank my primary supervisor, Dr Denis McKeown, for his endless support, enthusiasm and patience while guiding me throughout this PhD journey. Alongside your knowledge and expertise in memory research, I truly appreciate your encouragement throughout the years. My thanks also goes to my co-supervisor, Professor David Bunce, for his advice and guidance. This thesis was made possible thanks to the time and effort provided by my supervision team. A special thanks also goes to Dr Donna Lloyd and Dr Ed Sutherland for your useful advice and a friendly hello. Thank you to the Economic and Social Research Council and the White Rose University Consortium for providing the funding to conduct my work.

Thanks to all the participants who took part in my studies. I am especially thankful to the community centres and ageing groups who were so welcoming. For helping with data collection, thanks goes to Kinga Bugajska, Rose Wastling, Michelle Ames and Kayleigh Ramsey.

I am grateful for the opportunity to work alongside a brilliant team of postgraduate research students at the School of Psychology. A special thanks goes to my office buddies who made this experience so enjoyable. Thanks to Raginie Duara, Jihane Ghorayeb and Ruth Baxter for being a source of inspiration as the most hard working researchers in town. You may find at times you are down, troubled and nothing is going right. In these situations, it helps when someone is there to remind you that you can always turn to a career in tap dancing. To Sarah Wilding, Cheers for that! Outside of the PhD bubble, my thanks go to my best friend Jennifer Davis, for providing an escape to the big city with tea, cake and musical theatre on hand.

This thesis is dedicated to my family, for always believing in me even when I don't. To my dad, thank you for listening to my thesis ramblings and always being on hand to provide macaroni cheese when needed. To Our Rachel, thank you for being the best sister, always looking out for me and encouraging me to just keep on writing! Finally, thank you to the one and only Mama Smid. You have been a constant source of support and patience, putting up with my endless phone calls and panic. I can't express how much your help and encouragement has meant to me.

To end with a quote from my favourite wittertainment presenters, Mark Kermode and Simon Mayo. How do you write a doctoral thesis? You just WRITE IT! Producing this thesis has honestly been the biggest challenge I have ever undertaken. Please enjoy my best effort.

Abstract

Emerging evidence suggests that age-related declines in memory may reflect a failure in pattern separation, a process that is believed to occur in the hippocampal dentate gyrus/ region CA3 which reduces the overlap between two similar stimulus representations during memory encoding. Furthermore, this process of pattern separation may be indexed by a visual continuous recognition task (various names are popular in the literature, but BPS for behavioural pattern separation has perhaps been the most popular label) in which items are presented to observers in sequence and observers report for each whether it is novel, previously viewed (old), or whether it shares features with a previously viewed item (similar). Thus, worldwide a number of laboratories currently employ the task to infer not only aspects of human memory encoding and retrieval, but also make published claims as to the mapping of patterns of performance in the task (for example, accuracy in identifying novel items, or errors in reporting similar items as ones previously encountered) onto neural circuits and regions within hippocampus. Indeed, clinical claims have been made – notably that the behavioural task captures aspects of both normal neurocognitive ageing but also pathologies affecting human memory such as Alzheimer’s disease. This thesis reports a series of six experiments which in summary support certain constraints that may be recommended to future investigators; the work was inspired and guided by contemporary theories of human immediate memory. First, a critical variable – the temporal spacing of items within the task – was experimentally identified. The theoretical justification here was that post-item time permitted the memory trace of the items to be consolidated into memory (in one of the reported experiments post-item time was directly manipulated using the introduction of a secondary attentionally-demanding task). Ageing and performance on the task was also investigated, and errors in noticing novelty with age was an important aspect observed in performance. Further, one new finding reported here questions whether these memory deficits (failure in recognising change in ongoing stimulation) in older adults result from the interference caused by the encoding of intervening items, or from the degrading of a memory representation over an extended time delay. Finally, a recommendation is made to use stimuli that do not permit the participants to engage in verbal labelling and maintenance through verbal rehearsal – here a single ‘class’ or category of stimulus achieved this aim. In summary, this thesis reports important advances in understanding an influential experimental task claimed to tap hippocampal function – a function that may be termed ‘noticing the new’.

Contents

List of Tables

List of Figures

Abbreviations

η_p^2	partial eta squared, Effect size for ANOVA analysis
AD	Alzheimer’s disease
aMCI	Amnesic Mild Cognitive Impairment
ANOVA	Analysis of variance
BPC	Behavioural pattern completion
BPS	Behavioural Pattern Separation tasks
BPS-O	Behavioural Pattern Separation Task—Object Version
C	A criterion bias measure
CA1/CA3	Hippocampal sub regions of the brain
d	Cohen’s measure of effect size for comparing two sample means
d’	Discriminability, a measure of sensitivity in signal detection theory
DG	Dentate Gyrus, Hippocampal sub regions of the brain
fMRI	Functional magnetic resonance imaging
IE	Incidental encoding task
III	Inter-item interval
IRI	Item retention interval

M	Mean
Min	Minutes
MMN	mismatch negativity
MoCA	Montreal cognitive Assessment
Ms	Milliseconds
MST	Mnemonic similarity task
MTL	Medial temporal lobe
P	Probability
RAS	Recognition accuracy score
S	Seconds
SD	Standard Deviation
SIMPLE	Scale-independent memory, perception, and learning model
TBRS	Time based resource sharing model
Z	z score

Chapter 1 General Introduction

1.1 Overview of Thesis

In searching for the earliest notions of the memory trace in the psychology of memory, some would regard the Roman physician Galen as a good choice (Julião, Lo Presti, Perler & van der Eijk, 2016). According to Galen memory arises out of sensory perception leading to the formation in the brain of “impressions”. These impressions are able to capture the everyday, the events and objects encountered, through formed *traces* – the memories of the immediate past. Yet, for this formation of the memory traces to be robust, certain conditions must be met. The first condition is that the impressions must have clarity in and off themselves – be physically salient. The second is that attention to the everyday events must be applied by the perceiver. The third is that the body – the organism or person’s physiology and organs of sense – must be in a healthy state of reception. Thus, for instance, if an impression reaches the sense organs but these are not working properly, the impression might not pass the threshold of primary sensation, and therefore there is no transmission of information to a deeper level of storage. So people experiencing great stress, whose ‘souls’ are only dimly perceiving or they are inattentive, may fail to capture impressions. Therefore their memory traces will be dim too, and be maintained only weakly in memory.

This thesis is concerned with both the fate of the memory trace, the things that ensure it is secure in memory, and also some changes in the person – such as natural ageing – that may influence its formation and maintenance. More specifically, this thesis is concerned with behavioural measurements claimed to tap two memory processes termed ‘pattern separation’ and ‘pattern completion’ that indeed correspond rather closely to Galen’s notions of the coding of impressions and their re-awakening in retrieval. A surprisingly straightforward behavioural arrangement or task has in recent years been developed that, it is claimed, taps into these (largely hippocampal) fundamental processes within human memory; specifically, aspects of patterns of performance on this behavioural task are even claimed to map onto or model the complex relationships between discrete hippocampal regions (notably dentate gyrus and regions CA1 and CA3) performing reliable encoding and retrieval of items from memory. The contribution of the research studies reported in this thesis is to assess the utility and validity of the task, largely within the context of contemporary temporal- and interference-based accounts of forgetting over the short term. Chapter one will first outline the hippocampal processes, explain the behavioural task which is the focus of study, and then present six fundamental questions regarding the utility and validity of the task, namely: 1) it’s ability to chart clinical distinctions in cognitive impairment and memory; 2) it’s ability to reveal aspects of short-term memory encoding and retrieval (independently of verbal or longer-term associations within memory); 3) the ability of the task to chart aspects of human ageing and memory; 4) the role of short-term memory consolidation in performance on the task; 5) the attraction of manipulating the physical similarity of stimuli on the task; and 6) the role of stimulus novelty. The final Section of Introduction will explain the tasks, measures and analyses to be employed throughout. Six experimental chapters will then report tests directed at each of these questions. General discussion will assess the extent to which the questions have been successfully addressed, make some recommendations (and suggest some constraints), and propose some directions for further study.

1.2 Pattern separation and pattern completion as processes within hippocampus

Before setting out the types of behavioural measures under study – and especially the rationale for using the measures and the six questions of the validity of the task under study – the critical feature to keep in mind is that the focus is on the ability to discriminate between objects that have previously been encountered and objects that are novel; and further, to discriminate between two forms of novel object, namely ones bearing similarity to previously encountered objects and those fully novel. The terms that have been adopted within the literature to attempt to embrace such discriminatory processes are pattern separation and pattern completion. However, these terms are somewhat confusingly used in literature – they are used for specific (largely hippocampal) mechanisms or processes underlying memory encoding and retrieval – but also for behavioural indexes of the putative processes (where the valid application of the terms separation and completion are more debatable). The importance of human memory over the short term (from seconds to minutes, to hours) is realized when we consider the number of overlapping events we encounter daily. Despite the repetition of similar elements, our episodic memory system (Tulving & Markowitsch, 1998) can support mnemonic processes, rendering each episodes distinct in memory yet later accessible upon presentation of a subset of its original elements (Duncan, Sadanand, & Davachi., 2012). While this process of rendering elements distinct within the memory system appears effortless, it is underpinned by complex processes of encoding and retrieval. For example, when we encode visual items into memory, we need to ensure that elements that are similar, those sharing a number of features, are stored as distinct representations in order to reduce retrieval errors (McClelland, McNaughton, & O'Reilly, 1995). Evidence has converged from many sources, behavioural, anatomical, and neurophysiological as well as computational modelling, that the hippocampus supports the rapid encoding and retrieval via the processes of pattern separation and pattern completion (Hunsaker & Kesner, 2013).

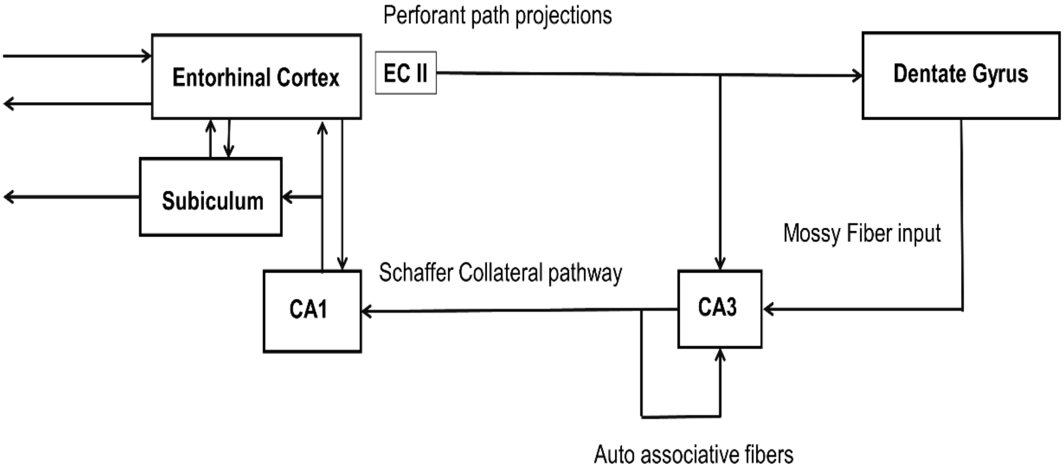

Figure 1-1. Schematic of entorhinal cortex and hippocampal subfield circuitry (Modified from Deuker, Doeller, Fell, & Axmacher, 2014).

Pattern separation is the term adopted for the process whereby similar or overlapping representations are stored as distinct orthogonal memories (Yassa & Stark, 2011). A slight detour into the underlying neuroscientific bases for this process is warranted here. This process of orthogonalization is believed to be essential for managing proactive interference generated by stored representations with shared features (Hunsaker & Kesner, 2013). Recent computational accounts of the hippocampus (McNaughton & Morris, 1987; O'Reilly & McClelland,1994;Treves & Rolls, 1992, 1994) propose that pattern separation is supported by the mossy fiber projections from the dentate gyrus (DG) to CA3 (Amaral & Witter, 1989). The high density of neurons in the DG, together with the very sparseness of the mossy fiber projections greatly decrease the probability of two different inputs to the DG, activating the same set of CA3 neurons (Myers & Scharfman, 2009, 2011). Casually, we might think of this as having a handful of rather similar pebbles and throwing them onto a wide sandy beach, so that the chance of any two lying side by side is unlikely. This is the critical neuroscientific idea underlying the idea of separation of features of representations, and underlies thinking about how a behavioural task might capture this ability when observers view objects.

The other process, pattern completion also has its underlying neuroscientific story. The great theoretician David Marr (1971) discussed the so-called ‘auto-associative’ properties of hippocampus. These are properties of a neural network (real or simulated) that capture relations between ‘input’ (for example a sensory pattern) and an ‘output’ (such as a category description for that pattern), using a form of learning known as Hebbian learning. This acts to adjust the connection strengths between the neurons (termed nodes in artificial modelling). The essential notion is that, following learning or training, such a network can retrieve a detailed representation based on noisy or incomplete input patterns. Therefore, it is believed that pattern completion is supported by the specialised circuity in the CA3 network. This enables retrieval of a memory item based on partial information for that representation. Pattern retrieval is achieved using the partial activations across a network (e.g. Rolls, 2007). Again, one component of the behavioural task used in this thesis is believed to capture aspects of this very process.

It is very important to mention that pattern separation and pattern completion are generally viewed as processes that complement one another (Yassa & Stark, 2011). The reader will be reminded of this fact throughout the present thesis, for it means that the behavioural outputs of the task used to assess the strength of each process (separation and completion) can never fully be regarded as independent. For example, when presented with a stimulus sharing similar features to those of a previously stored item, the ability of the hippocampus to accurately retrieve the original stored item may rely on the fidelity of the encoded representation. That is – and maybe rather obviously – retrieval depends on the initial pattern separation processes, which enabled a distinct representation to be formed or laid down. Furthermore, based on their computational definitions, Hunsaker and Kesner (2013) highlight that pattern separation and pattern completion are somewhat ‘at odds’ with each other. Pattern separation reduces overlap between representations but in doing so, it hinders the pattern completion processes which must use overlapping patterns of activation as a cue for accurate retrieval. Therefore, successful encoding and retrieval may be dependent on the balance between the two processes. Of course it is not controversial to state that retrieval and encoding of information are interdependent – but in the present work it will be kept in mind (and it will instruct present concerns that the behavioural task is inherently flawed as it’s measures are interdependent rather than pure – see General Discussion).

It is also worth noting that recent reports of pattern separation and pattern completion have suggested that an imbalance between the two processes can be attributed to mnemonic similarity of the memory representations (e.g. Stark, Stevenson, Wu, Rutledge, & Stark, 2015; Yassa, Matterfeld, Stark, & Stark, 2011) or damage to the DG/CA3 region associated with neurocognitive ageing (Stark, Yassa, Lacy & Stark, 2013; Stark et al., 2015; Wilson, Gallagher, Eichenbaum & Tanila, 2006); or specific memory impairments such as mild cognitive impairment and Alzhimer’s disease (Yassa et al., 2010). As explained in the next section, one exciting focus of current research into these processes has been to establish behavioural measures that complement the growing body of neuroscientific evidence, and such measures may be used within clinical practice (for example, in discriminating between forms of neurocognitive impairment, or charting effects on memory of normal ageing).

1.3 The three-response task

For visual object recognition, research activity has been directed at uncovering behavioural indexes of pattern separation and pattern completion using continuous recognition paradigms (Hunsaker & Kesner, 2013). In these tasks people are presented with a series of discrete computer-presented pictures of everyday objects (for example, a teacup, a tree, a vehicle). For each picture in the series, the task is to report whether it is one observed before: ‘old’, not observed before in the series: ‘new’, or one slightly changed: ‘similar’ (Kirwan & Stark, 2007). Typically the old and similar stimulus items share features, and can vary in the number of shared features. Within the series, old and similar items are typically separated by several intervening items. Now, the interest is in the patterns of responding across the three response types (new, old, similar), and in the patterns of errors made too. Authors such as Kesner have made a persuasive case for indexing ‘separation’ in terms of the relative numbers of intervening items, on the assumption that some form of interference occurs as successive items are encoded into immediate memory. Naturally, the encoding of the ‘similar’ stimulus item is assumed here to form a representation sharing a large number of features with the encoded ‘old stimulus’ representation. The suggestion by Stark and colleagues is that, in order to separate and reliably form distinct representations, with minimal ‘overlap’, a form of encoding separation must occur within hippocampus. Yet, conversely, the assumption is that the associations between the pair of similar items are sufficient to engage pattern completion under some circumstances. As discussed below, one such circumstance may be age-related reliability of initial encoding.

Before considering the task in more detail, a very important point (and one informing the experiments reported) is that, although ‘old’ and ‘similar’ items may be perceptually distinct they may nevertheless share a common category or verbal label

(Stark et al., 2015) to the extent they represent the same concrete object (Pidgeon & Morcom, 2013). This – as explored below – can present as a major difficulty in interpreting data arising in the three-response task. If, for example, a participant in a study using different types of ‘cup’ stimulus is relying on verbal labels (exclusively or predominantly) they are likely to report ‘old’ – they are in other words likely to miss on similar or new trial presentations of cups. So it is (and this is a point noted in the paper by Hunsaker & Kesner, 2013 referred to above; also by Liu, Gould, Coulson, Ward & Howard, 2016) desirable to encourage participants to focus on physical features of presented items – and also of course to avoid using items that lend themselves to easy verbal labelling. As discussed below, perhaps the major contribution of the present reported studies is to use stimuli that share a single category label (the one here is ‘doors’) and so make it difficult or impossible for participants in the task to rely on category labels – that is, they must search for distinguishing physical visual features.

There are some variants, but the basic three-response task will be described first. Building upon a visual recognition paradigm of Koutstaal and Schacter (1997), Kirwan and Stark (2007) introduced a continuous recognition task to attempt to differentiate between the pattern separation and pattern completion processes outlined in the preceding section. Participants view a computer screen and a series of pictures of everyday objects are presented in succession (e.g. a wheelbarrow, then an apple, then a table) and for each the required response is to report by button press whether the item is “new” (an item was presented for the first time in the series), “old” (an exact repetition of an item encountered previously in the series) or “similar” (an item presented for the first time, which shared similar visual features with a previously viewed item). In an early report using this simple task, Kirwan and Stark (2007) presented items on screen for 2500ms with a brief half second inter-trial interval, and Figure 1.2 illustrates their procedure.

One key reported outcome measure for this ‘BPS’ task is the proportion of responses (reporting new, old or, correctly, similar) to the similar items. That is, studies employing the BPS task use the proportion of similar responses to similar items as a behavioural outcome of a bias towards pattern separation processes (e.g. Anderson, James, & Kirwan, 2017; Doxey & Kirwan, 2015; Duncan et al., 2012; Kirwan & Stark, 2007; Holden, Toner, Pirogovsky, Kirwan, & Gilbert, 2013; Toner, Pirogovsky, Kirwan, & Gilbert, 2009; Yassa et al., 2011). Whether or not such an assumption is justified will be a corner-stone of this thesis discussion, but for now the literature will be taken at face value. So, in order to control for response bias, a standardised index of behavioural pattern separation is calculated called the Behavioural Pattern Separation score (BPS; see Table 1.1 below) – based on the proportion of similar responses to similar items (termed pattern separation rate) minus new responses to similar items (termed similar bias rate) (e.g., Ally, Hussey, Ko, & Molitor, 2013; Anderson et al., 2017; Déry, Goldstein & Becker, 2015; Déry et al., 2013; Shelton & Kirwan, 2013; Yassa et al., 2011). Please note that whilst this calculation is general employed, some authors have proposed an alternative to ‘similar bias rate’, using instead a form of false alarm rate of misidentification of similar items as old (sometimes referred to as ‘pattern completion bias’). The final section in Introduction considers these matters further, and General Discussion will further question the validity of derived scores.

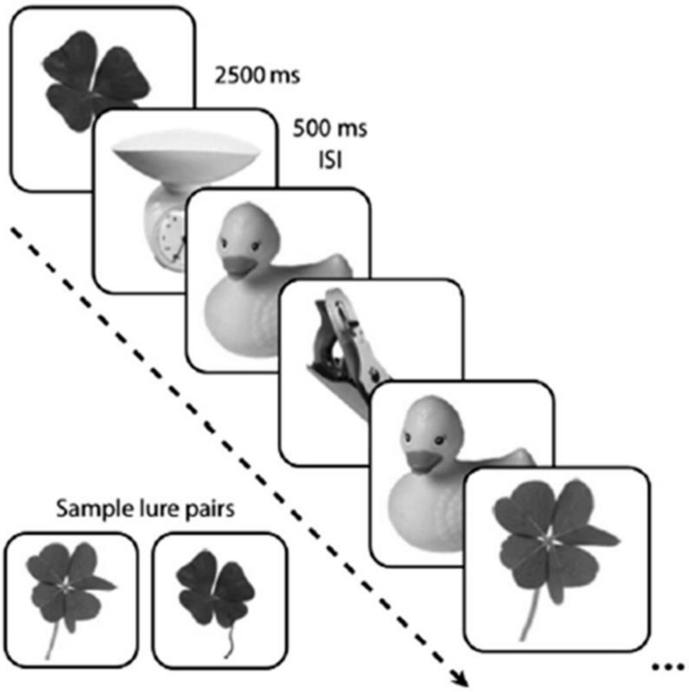

Figure 1-2. Schematic of the Behavioural Pattern Separation Task (BPS; Kirwan & Stark, 2007). Each item is presented on screen for 2500ms with an inter-item interval of 500ms. Each experimental block consists of 108 trials, with 44 new items, 16 old items (an item repeated twice in the sequence) and 16 similar items (consisting of similar item pairs which share similar visual features). Items are presented in a pseudo random order with repeated and similar item pairs separated on average by 30 trials (range = 1-105 trials). Figure taken from Yassa et al. (2010).

Now consider how pattern completion is assessed in the task. Certain computational models (Hunsaker & Kesner, 2013) define pattern completion as the retrieval of a previously stored item in response to a partial or downgraded cue – and of course this is what the similar stimulus item is. It is (generally) believed that completion is captured in the BPS task by the proportion of similar items incorrectly identified as old (Kirwan & Stark, 2007; Toner et al., 2009; Yassa et al., 2010). The reasoning behind this is that when a similar item is identified as old, it has acted as a partial cue for the recall of a prior item (Toner et al., 2009). Here we see an inter-play between noticing change in a stimulus (it is not exactly the one seen earlier) and making a decision of the form, “this is very much the same as the item seen earlier”. Presumably an observer has to both identify old features and also novel features and assign some weighting between them – there may in some instances be a weighting to report ‘old’ and in other cases a weighting to report ‘similar’; in contrast a failure to notice the old features will likely result in a “new” response. In summary, for similar items the task for an observer in the BPS task is to accurately identify that while the present item is familiar, being similar to items just encountered in the session, it is not precisely the same as the original item (Kirwan & Stark, 2007). As shown in Table 1.1, the Behavioural Pattern Completion score is calculated as the proportion of old responses to similar items (termed the pattern completion rate) minus the proportion of old responses to new items (termed the false alarm rate). Again, please note there are some who question the ‘false alarm rate’ and instead propose that old responses to similar items would be a more valid choice; the final section in Introduction revisits this matter. Below the question of whether these derived scores are valid, or indeed the are most attractive forms of derived score within the data, is considered.

Table 1-1.

Definitions of the response possibilities for the new, old and similar stimuli; and explanation of the two main derived measures BPS and BPC.

Stimulus	Response		
	New	Old	Similar
New	Correct rejection Rate	False alarm rate	Similar bias rate
Old	Miss rate	Hit rate	Incorrect
Similar	Incorrect	Pattern completion rate	Pattern separation rate

Behavioural pattern separation (BPS) = (pattern separation rate) – (similar bias rate)
Behavioural pattern completion (BPC) = (pattern completion rate) – (false alarm rate)

One way the three-response task has been developed is a form known as the incidental encoding task (IE; Bakker, Kirwan, Miller & Stark, 2008). Here, in a first phase an observer passively views a continuous sequence of new, old and similar items with no overt memory response requested; instead, typically, some form of response is demanded to ensure participants are paying attention to the items, such as stating whether the picture is of an indoors or outdoors object (e.g. Azab, Stark, & Stark, 2014); a toy or not a toy (e.g. Motley & Kirwan, 2012); pleasant or unpleasant (e.g. Huffman & Stark, 2014). The key feature is that during this viewing phase, participants are unaware that the task is one testing recall. One might assume the task is tapping a passive form of encoding of the pictures – or at least that they will not be engaging in some form of active ‘rehearsal’ or ‘refreshing’ (see below for an explanation of this idea) of the stimuli. Then in a second phase (see Figure 1.3), the participants are invited to complete the standard form of the three-response recognition task, where a sub-set of the stimulus items had been presented in the first phase. The rationale is that this two-phase task will circumvent the undesirable situation where participants are actively trying to maintain items in short-term working memory; such active rehearsal would, it is believed, confound attempts to uncover ‘normal’ encoding and retrieval processes.

It is worth making a note here that in certain studies, throughout this task, participants are undergoing high-resolution functional magnetic resonance (fMRI) scanning of the sub-regions of the medial temporal lobe (MTL). The intention is to see whether pattern separation/completion processes might be identified according to the level of activity recorded in fMRI response to each item (new, old and similar) in the three-response recognition second phase of the task (Azab et al., 2014; Bakker et al., 2008; Huffman & Stark, 2014; Lacy, Yassa, Stark & Muftuler, 2011; Motley & Kirwan, 2012). The expectations in such studies rely on particular models or schematics of distinct regions and circuits within hippocampus in encoding new items and recognising old ones (or identifying that new items share a lot of features with ones previously encountered).

These broadly similar tasks – for the differences between them are small - use a continuous series of items such as the computer-presented pictures described in the previous section, including repeated and highly similar visual stimuli, to provide a behavioural index of recognition memory, pattern separation and pattern completion. These tasks vary slightly according to the way the items are organized in the sequence presented to participants, as well as the types of behavioural responses recorded (Hunsaker & Kesner, 2013; Liu et al., 2016). In order to fully outline the current behavioural evidence for pattern separation and pattern completion, it is therefore necessary to outline each of these types of task in order to define the key outcome measures, recorded as a proxy of these two processes of encoding (pattern separation) and retrieval (retrieval). For now we will take at face value some of the claims made for the efficacy of the task – but become more critical as the present series of experiments are introduced.

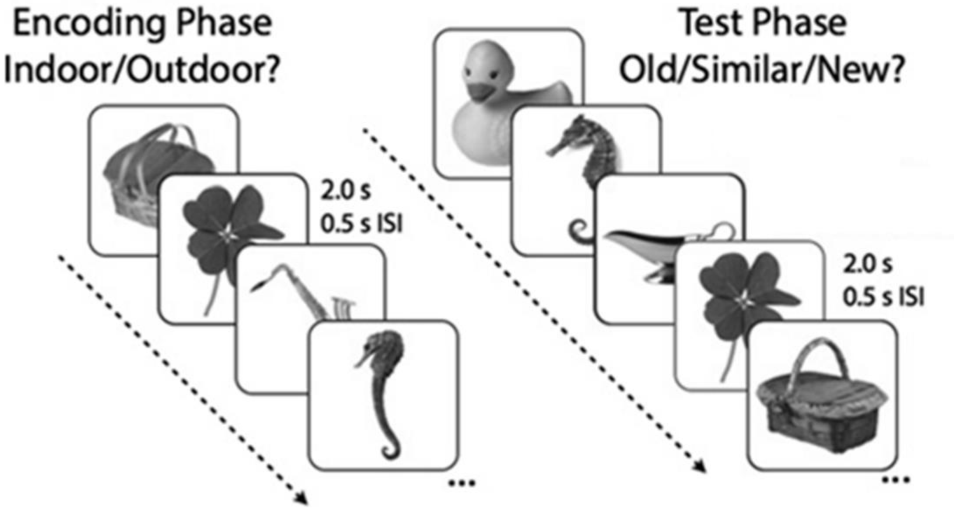

Figure 1-3. Schematic of the Behavioural Pattern Separation task-object version (BPS-O). Figure taken from Stark, Yassa, Lacy & Stark (2013).

A variant of the above two-phase task with an incidental encoding and recognition memory phase using a database of visual object pictures provided by the authors, termed variously the BPS-O Task (O for objects, referring to the stimuli which are pictures of real-world objects) and the Mnemonic Similarity Task (MST) (Stark et al., 2015). These authors not only provide the stimuli but provide an 'on-line' test useful for investigators or clinicians (in this way many studies in the literature – perhaps most – apply the task, assumptions intact). Clearly, though, one needs to keep a careful eye on the derived measures (such as the BPS and BPC scores described above) – they may differ across studies (see General Discussion).

1.4 The question of clinical measures of cognitive impairment

One exciting development – and one inspiring the start of the present research effort – is that some claim that the BPS task is a neuropsychological test capable of identifying the early behavioural markers of cognitive impairment related to neuropathological conditions, such as amnesic mild cognitive impairment (aMCI; Yassa et al., 2010) and Alzheimer’s disease (AD; Stark et al., 2013). Using the BPS task, Yassa et al. (2010) found that patients with aMCI showed a reduction in pattern separation score. In comparison to healthy age-matched adults, patients with aMCI showed an increase in the proportion of similar items incorrectly identified as old. Further, in a recent review (Ewers, Sperling, Klunk, Weiner, & Hampel, 2011), patients with late aMCI and early AD demonstrated reduced overall hippocampal activity. Specific to pattern separation processes, elevated neural activity in the CA3 region is claimed to result in an inability to form new memories (Wilson, Ikonen, Gallagher, Eichenbaum, & Tanila, 2005); and across two randomised control trials, Bakker, Albert, Krauss, Speck, & Gallagher (2015) found that increased activity in the DG/CA3 correlated with lower performance overall on the BPS. Yet, following treatment of a low dose of antiepileptic levetiracetam, there was a significant increase in performance on the BPS task, consistent with reduced hyperactivity in DG/CA3. These important findings have encouraged clinicians to adopt the on-line version of the BPS for assessing and distinguishing between neurocognitive conditions.

Despite such encouraging relations between performance on this simple behavioural task and neurocognitive impairment, it remains unclear how patterns of performance (such as certain derived measures explained in Introduction below) might differentiate between aMCI, AD and levels of cognitive performance (such as short-term memory recall) in a normal ageing population. One suggestion (Ally et al., 2013) has been to ‘load memory’ in the BPS test by manipulating the number of intervening items between 'old' and 'similar' item presentations – presumably such a manipulation would introduce not only a greater number of observed items overall in the intervening sequence, but also introduce confusion between items, especially between ones falling within the same category (domestic utensils, animals, vehicles etc). Ally et al.

measured pattern separation performance in healthy older adults, aMCI, and AD groups, at differing numbers of intervening items (4, 12 or 40), a variable the authors termed 'lag'. They report that pattern separation performance was lowest for AD patients, middling in the aMCI group and highest in the normal healthy ageing group. More interesting is how the patterns of performance changed with increasing lag. First, the AD showed flat, poor or chance performance across increasing lag; second, the normal group performed reasonably well at each lag. In marked contrast, pattern separation performance in the aMCI group actually decreased with increasing lag. Ally et al. proposed that this distinct pattern of BPS performance in aMCI is evidence of rapidly degrading visual representations in immediate visual memory, of the order of seconds.

So a key starting focus of the present thesis was to study how recognition memory changes with increasing ‘lag’. As discussed below, there was good reason in the Leeds laboratory to believe that the BPS lag manipulation did more than simply introduce additional items into the gap between an ‘old’ and a ‘similar’ item. Certainly, in studies of healthy adults (or at least people without cognitive impairment), there is consensus that reliable item encoding within immediate memory (this term is adopted here and throughout to refer to memory over the short term of seconds and minutes, but distinct from any acceptance of the familiar ‘short-term’ within the familiar ‘working memory’ models of Baddeley and colleagues; cf. Baddeley & Hitch, 1974) – which can be described as pattern separation - is supported by circuits within hippocampus. It will be worthwhile briefly outlining certain developments using the types of old, new, similar tasks indexing pattern separation in recent studies. High resolution fMRI studies – no attempt being made to fully decipher the statistics of the reported fMRI outcomes in these studies – reveal that pattern separation versus pattern completion are region-, or at least circuit-specific within hippocampus. Broadly, it is fair to say that changes in activity in the hippocampus have been found to predict behavioural responses to similar items (Kirwan & Stark, 2007; Kirwan et al., 2012; Reagh & Yassa, 2014; Yassa, Lacy et al., 2011;Yassa, Matterfield et al., 2011). For example, in Kirwan and Stark (2007), participants completed the BPS three-response task. While accuracy was high for the identification of new and old items, responses to similar items were mostly divided between correct responses (similar) and the incorrect response ‘old’ (that is, few incorrect ‘new’ response to the similar items were observed). This pattern of responses is often observed in the behavioural separation task (Stark et al., 2013), with many interpreting this as a bias towards pattern completion (Yassa, Matterfield et al., 2011; Kirwan & Stark, 2007). An interesting confirming observation in the scanning data is that neural activity in the hippocampus during the first presentation of an item predicts later behavioural successes; pattern of activity apparently distinguishes between incorrect old responses and correct similar responses to similar items. Such a report of course greatly strengthens the argument that the behavioural task is indeed capturing something of the underlying circuitry-generated memory mechanisms. Perhaps, given the somewhat ‘broad-brush’ resolution of fMRI (the hippocampal circuits and distinct loci are after all highly localised and not extensive), it is fair to say that the picture within the neuroscience related to the three-response task is encouraging if hazy. Therefore, it is critical that we arrive at a better understanding of the BPS task itself – before simply accepting that it allows clinical assessment of ageing and cognitively impaired populations. The first question then is, what is the ‘lag’ manipulation of Ally et al. described above really telling us?

1.5 The question of developing valid stimuli

Every psychologist must ask of the task they choose, does it measure what it is supposed to measure? This question is perhaps the one uppermost in the mind in the design of the present set of experimental procedures. If we accept that the functional properties of hippocampal circuits are broadly correct – namely, that some parts force separable representations of experienced objects or events (minimizing interference within encoding through pattern separation), whilst some parts allow retrieval of a memory trace of those representations even when only some part of the object is available within the input pattern (allowing pattern completion) – the question of how such functional properties may be captured in simple response decisions in participants is challenging. This is not the same as saying a response decision can be mapped directly onto activity at the cellular level – it would be naïve to make too direct a claim even for broad properties of neuronal collections or modules – but what might be claimed is that the task does reveal the same types of process. The usual term is functional equivalence. This idea is accepted in many areas of psychology. An example is the frequency encoding by the peripheral auditory system studied by McKeown and colleagues (e.g. McKeown & Patterson, 1996; Patterson et al., 1991). The peripheral auditory system (cochlea and some higher level modules) acts like a set of filters using a form of mathematics called Fourier analysis – allotting sound information to different channels – but no-one is claiming that the inner area is running a mathematical operation.

Yet, at the least the three-response task should be capturing the main properties of the computational processes believed to be performed in hippocampal system. And at face value certainly the task does demand discriminations between stimuli as well as generalisations between stimuli, the two essential properties. Thankfully some authors (e.g. Lui et al., 2016) have carefully considered and proposed a number of ways in which we should assess whether the task is valid or not, beyond simply appearing so at face value. Three essential properties may be identified which have informed the development of the present form of BPS task. First, the test should be tapping into the immediate past rather than longer-term storage in tapping pattern separation – therefore the stimuli should be novel not ones leading to retrieval of longer-term associations (which would of course be invoking pattern completion). So, for visual stimuli the standard use of pictures (even the somewhat cartoon-like ones used generally, and in the BPS-O on-line set of stimuli) of everyday familiar objects is far from ideal. It would be better to use never seen before or abstract objects or images. Such novel images offer an important additional advantage – they do not promote verbal labelling, and as suggested by McKeown and Mercer (2012), studies of short-term memory would do well to avoid verbal or category labelling since it allows participants to use a verbal-label memorisation strategy (so the investigator of memory encoding and retrieval is not tapping anything other than the ability to rehearse and recall words).

A second essential property of the three-response task is that the part of the task tapping pattern separation should properly be tapping encoding, whilst the part tapping pattern completion should be tapping retrieval. This desirable distinction may be impossible to achieve. If I do not accurately encode an object or feature of an object my retrieval will probably fail; or at least, retrieval will be impaired. Nevertheless we should attempt to get a clear picture of how we derive certain statements about responding in the task. The main point here is that in our analysis of responses across new, old and similar types we should be restrained, recognising the inter-dependence of the ‘rate’ measures.

1.6 The question of ageing

Renewed interest in developing accurate measures of behavioural pattern separation and pattern completion has in part been driven by an attempt to better understand older adults’ memory and forgetting. Indeed, a growing body of evidence suggests that older adults exhibit a deficit in encoding new memories so that they are distinct from previously stored items. In other words, they show impaired pattern separation (Carr, Castel & Knowlton, 2015; Holden et al., 2013). For example, research reveals that visual recognition impairments in older adults may be due to an impaired ability to identify stimulus novelty (Yassa & Stark, 2008). As we age, impaired recognition of everyday objects increases (Norman & Schacter, 1997), which can result from novel items being viewed as though they had been previously seen (Koutstaal & Schacter, 1997). Yeung, Ryan, Cowell and Barense (2013) presented older adults with a series of everyday objects in an initial study phase. They then recorded their eye tracking behaviour whilst they viewed some of the objects from the study phase amongst new objects, which shared either high or low similarity with previously viewed objects. Commonly, mean eye fixation is greater for the exploration of a novel object (Henderson & Hollingworth, 2003). Yet Yeung et al. found that mean eye fixations in their older adults did not differ across old repeated items and items sharing high similarity; apparently new items were falsely viewed as old. The authors noted that this outcome pattern may have reflected either impoverished encoding in the study phase (Molitor, Ko, Hussey, & Aly, 2014) or simply failure to identify novel visual features. This is clearly a fundamental question with respect to human memory, forgetting and ageing and understanding – and unravelling - the contribution of the Yeung et al. data will play a central role in the present research endeavour. Since ageing plays such a role it is necessary to take a diversion into the human ageing literature, although maintaining the focus on pattern separation and pattern completion as much as possible.

Impaired performance on BPS. Evidence of impaired object recognition in healthy older people has been documented using the BPS and BPS-O tasks (Bakker et al., 2012; Bennett & Stark, 2016; Brickman et al., 2014; Doxey & Kirwan, 2015; Holden et al, 2013; Reagh & Yassa, 2014; Stark et al, 2015; Stark et al., 2013; Toner et al., 2009;

Yassa, Matterfield et al., 2011; Yassa et al., 2010). Age related impairments in encoding has long been attributed to impaired mnemonic discrimination due to a decline in the function of the DG (Brickman et al., 2014; Deuker et al., 2013). In their computational model of neurocognitive aging, Wilson et al. (2006) proposed that this decline in function in the DG may be attributed to the reduction in synaptic connection from the EC via perforant pathway (Geinisman, deToledo-Morrell, Morrell, Persina, & Rossi, 1992; Smith, Adams, Gallagher, Morrison & Rapp, 2000). In turn a loss of synaptic connection in DG to CA3 creates an imbalance between pattern separation and pattern completion (Hasselmo & Schnell, 1994; Hasselmo & Wyble, 1997), which may bias the system towards pattern completion (Wilson et al., 2003; Yassa et al., 2011). In other words, in the three-response task we will observe similar items being reported as old (a failure of pattern separation – or a bias towards pattern completion).

Behavioural measures of pattern separation have identified impaired object recognition in in healthy older adults, by comparing performance between two distinct age groups, younger adults, typically aged 18-30 years and healthy older adults, aged over 60 years (Doxley & Kirwan, 2015; Reagh & Yassa, 2014; Stark et al., 2015). Interestingly, performance between the two groups typically is not seen to differ for the measures of recognition memory, captured by the accurate identification of new items and old items. However, younger adults are observed to outperform older adults for the correct identification of similar items (Stark et al., 2015), indicating reduced pattern separation ability. Toner et al. (2009) therefore suggest that the fact older adults are more likely to incorrectly identified similar items as old indicates a bias towards pattern completion.

Older people may be more prone to proactive interference. Stark et al. (2013) demonstrated that behavioural pattern separation scores in the BPS gradually decline across the lifespan. Based on this, it could be suggested that diminished ability to recognize the novel features of objects results from neurocognitive ageing (i.e. Wilson et al., 2006), where in older adults the formation of new memories is hindered by interference from prior memories. That is, maybe older adults are more prone to proactive interference. When the stimulus similarity between item pairs in the task is increased, performance diverges more obviously between older and younger adults. For low similarity pairs, pattern separation scores are equal for younger and older adults but when similarity is increased, older adults' ability to discriminate between item pairs decreases (Yassa, Lacy et al., 2011; Yassa, Mattfeld et al., 2011). Toner et al. (2009) argued that age related changes to the hippocampus in older adults may result in inefficient pattern separation, rather than a recognition memory deficit per se.

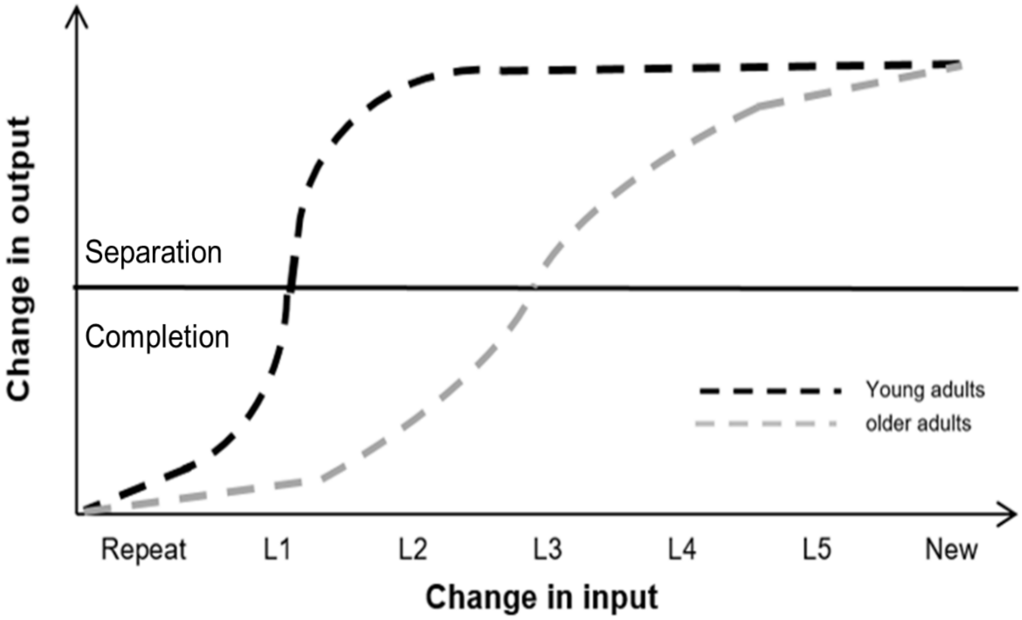

Figure 1-4. Demonstration of representational rigidity in older adults (adapted from Yassa, Mattfeld, Stark & Stark, 2011). Hypothetical data represents activity in DG/CA3 as a function of change in input. L1 represents a small change (highly similar) input and L5 represents large change (low similarity). The horizontal line represents the switch between pattern separation and pattern completion. In order to engage in pattern separation, older adults require a greater dissimilarity between the input and output.

Structural changes in hippocampus. Evidence suggests that the association between age and performance on behavioural pattern separation tasks is underpinned by the structure and activity of the dentate gyrus and the CA3 (Yassa, Lacy et al., 2011). Again, the claim is that age related changes to perforant path and activity in DG/CA3 will bias against pattern separation, contributing the impaired mnemonic discrimination (Yassa, Mattfeld et al., 2011). Using the incidental encoding version of the task and using fMRI scanning, Yassa, Lacy et al., (2011) investigated changes in activity in the DG/CA3 whilst manipulating item similarity. For dissimilar items, there was no significant differences between activity in the DG/CA3 between older and younger adults; but in older adults, DG/CA3 activity declined as similarity between similar items increases. In an additional recognition task, the ability to accurately identify similar items declined with increasing stimulus similarity. In older adults, this change in response in the DG/CA3 region is thought to be reflective of a ‘representational rigidity’. Therefore in order for older adults to discriminate between similar memory representations, a larger change in input between item pairs is needed to promote a bias towards efficient encoding (Yassa, Lacy et al., 2011)

Reliance on ‘gist’ memory. Unfortunately a simple ‘familiarity judgement’ in BPS responding may be insufficient to reveal the imprecision of stimulus encoding presumed with age. Certainly, false alarms (reporting new items as old) are notable in older age samples in the sorts of task used here (Edmonds, Glisky, Bartlett, & Rapcsak, 2012). Recently, Devitt and Schacter (2016) provided a broad overview of the effect of aging on the medial temporal lobes and the prefrontal cortex (PFC) to identify the cognitive processes underpinning such response errors. Firstly, inefficient pattern separation during initial encoding of a memory item may lead to a reliance on gist memory which is prone to error (Morcom, 2015). Secondly, of course, the problem may lie at the retrieval stage; Koen and Yonelinas (2014) have argued that, even when a memory item is successfully encoded, older adults demonstrate poor recollection strategies hindering the comparison of incoming stimuli to the previously stored memory items (also suggested by Pitarque, Meléndez, Sales, Mayordomo, Escudero et al., 2016).

However, consider the notion of a ‘recall to reject’ strategy outlined in some reports. This is the idea that to accurately identify similar items, observers must first recall the stored memory of the initial item (engaging pattern completion processes; Kirwan & Stark, 2007; Morcom, 2015). Therefore, the behavioural index of pattern separation is underpinned by accurate encoding of similar items as well as retrieval of stored items, potentially engaging pattern completion processes (Yassa et al., 2010). Presumably, under conditions of high similarity between successive items, such a strategy will depend upon the precision with which items are encoded or maintained. Within the visual working memory literature, there is growing evidence that memory precision declines with age (Peich, Husain, & Bays, 2013; Pertzow, Heider, Liang, & Husain, 2015; Souza, 2016). Souza (2016) demonstrated that even if only a single feature of an item is required to be retained in memory, older adults still show lower memory precision. Souza conducted a continuous colour reproduction task in which participants were required to identify the colour of a cued circle from a previously studied memory display. Older adults demonstrated not only a higher rate of recall error but reduced memory precision, as indicated by the discrepancy between the colour they selected on a colour wheel and the true colour of the target item. Therefore, a recall to reject strategy will be prone to failure.

1.7 The question of time and memory consolidation

Perhaps the most worrying feature of the three-response behavioural assay of PS and PC is that there is often an unquestioning acceptance of the basic tenets, often uninformed by current theory on short-term memory and forgetting. In this section the task is reconsidered in the light of contemporary work on time-based forgetting. Contemporary theoretical explanations of maintenance and forgetting in human memory have generally contrasted mechanisms of time-based decay on the one hand and interference-based corruption of the memory trace on the other (Barrouillet & Camos, 2014); although it is fair to say a growing consensus is that both decay and interference play a role in forgetting, with slow time-based decay processes acting as a form of 'eraser' for redundant memory content (Altmann & Gray, 2002; Altmann & Schunn, 2012). Decay of the trace since encoding in this view is an adaptive process which removes residual representations, and by doing so reduces proactive interference (Hardt, Nader, & Nadel, 2013; Mercer & Duffy, 2015). Decay also plays an important role in an influential current model of attention-based memory, the time-based resource sharing model (TBRS; Barrouillet, Bernardin, & Camos, 2004; Barrouillet, Bernardin, Portrat, Vergauwe, & Camos, 2007; Vergauwe, Dewaele, Langerock, & Barrouillet, 2012; Vergauwe, Hartstra, Barrouillet, & Brass, 2015). Here, though, it is not time since encoding itself that produces a degrading of the memory trace, but rather the proportion of that temporal interval where attention is otherwise captured and therefore is unable – on the assumption that attention is a limited resource - to maintain the memory trace. Yet others have argued that time-based decay plays no role in forgetting, but rather loss of information in memory is due to event-related interference (e.g. Lewandowsky, Oberauer, & Brown, 2009), or to the overwriting of features of the memoranda by distracting material presented post-encoding (Oberauer & Kliegl, 2006). Time in such conceptions does not provide an opportunity for decay but rather beneficially isolates memoranda within psychological space and protects them from being confused with other events entering into memory.

Time based forgetting in short term memory is well documented both in the laboratory at Leeds (McKeown, Holt, Delvenne, Smith, & Griffiths, 2014; McKeown & Mercer, 2012; Mercer & McKeown, 2014) and elsewhere (Zhang & Luck, 2009), yet the underlying mechanisms of such decline in memory over time are fiercely debated (Altmann & Schunn, 2012; Ricker, Vergauwe & Cowan, 2016). The research question is one familiar in the early twentieth century accounts of forgetting: does forgetting occur because of decay occurring as time is extended, or does it occur because of the disruptive influences of other events occurring as time elapses?

One account of such time-based influences appeals to the temporal distinctiveness of items within their spatial and temporal context (Ecker, Brown, & Lewandowsky, 2015; Brown, Neath, & Chater, 2007). It is thought that increasing the inter-item interval in recall task reduces the confusion between memory items on a trial by trial basis (Ecker, Tay, & Brown, 2015; Mercer, 2014). Temporal distinctiveness accounts capture this idea of a critical inter-item interval between items for recall by proposing that that successful memory retrieval is determined by the ratio between the inter-item intervals between successive to be remembered items and the retention interval between initial encoding of an item and subsequent retrieval (Ecker, Tay, et al., 2015; Souza & Oberauer, 2014). It has been argued that this ratio is critical for successful memory recall because when memory items 'retreat' into the past they become compressed (Grange & Cross, 2015) and by reducing the inter-item interval, confusion may arise between those items within their spatial and temporal context (Ecker, Brown, et al., 2015; Ecker & Lewandowsky, 2012).

One of the strengths of the temporal distinctiveness account is that specific predictions may be made for a given schedule of presentation of memoranda based on relative time. A recent study by Souza and Oberauer (2014) manipulated the ratio between inter trial interval (ITI) and memory retention interval (RI), to investigate the effect of temporal distinctiveness separate from simple decay. In their continuous colour recall task (similar to one influentially used by Zhang & Luck, 2008), participants were presented with a visual item consisting of six coloured circles. Following a brief delay (RI), a colour wheel was presented with a cue to the location of one of the coloured circles. Participants were required to recall the colour of the cued item. Based on manipulations of the RI and ITI, Souza and Oberauer were able to estimate the distinctiveness of each memory display. Using these calculations, they found that performance did indeed vary as predicted by the temporal distinctiveness ratio.

Now, Souza and Oberauer (2014) applied a simple ratio of retention interval (current trial between memory display and probe) and the time elapsed since the presentation of the memory display on the previous trial. Could such calculations apply to the continuous recognition BPS task, so as to account for some aspects of performance? Consider a probe item that repeats an 'old' item from five back trials in the BPS task. The retention interval is not within the 'current trial' in this case but is the time elapsed since the old item occurred. In comparison to Souza and Oberauer, the ‘time elapsing since the prior trials item’ is, in the BPS task, the inter-item interval. Therefore, whether considering the distinctiveness of items either in memory or simply within the current trial, it is obvious that calculations in the BPS task are confounded by the number as well as the spacing of the intervening items between memorandum and its re-occurrence on the current trial. So, despite the growing body of evidence in support of the temporal distinctiveness accounts (Ecker, Brown, et al., 2015; Ecker, Tay, et al., 2015; but see Ricker & Cowan, 2014 and Ricker, Spiegel, & Cowan, 2014 for conflicting results), it remains difficult to precisely predict the outcomes in the BPS task. Do other contemporary memory accounts offer better promise?

One form of item interference account in immediate memory (Oberauer & Lin, 2017) proposes that the continuous encoding of visual items results in the super imposition of new information onto items already presented. This might occur in any memory test where for example an intervening or interfering item occurs between a test item and its subsequent recall. Consider for a moment the idea that an item in memory is stored in terms of its content and context. If access to the stored item is cue based retrieval, interference occurs when the cue is not unique to the target item but instead results in the retrieval of the target item as well as non-target information (Oberauer, Lewandowsky, Farrell, Jarrold, & Greaves, 2012). In other words what occurs is distortion of the memory representation (Oberauer & Lewandowsky, 2008). However, it has been suggested that the degree of distortion depends on the encoding strength of the interfering item (Awh, Barton, & Vogel,2007; Awh & Vogel, 2008).

Perhaps the key factor governing encoding strength is stimulus novelty (also see section on novelty in this Chapter below). Introducing perceived novelty of intervening items in memory tests has been identified as a source of retroactive interference, both in verbal stimuli (Ecker, Lewandowsky, & Oberauer, 2014; Lewandowsky & Oberauer, 2009; Oberauer, Lewandowsky et al., 2012) and in visual stimuli (Chen, Cook & Wager, 2015; Hashimoto et al., 2012; Howco & Lepage, 2014; Moran & Goshen-Gottstein, 2015). In a continuous sequence of items as in BPS task, each item encountered may be assigned an encoding strength, which is based on the perceived novelty of that item. If a stimulus is repeated, the encoding strength assigned to that item is reduced. This in turn decreases the level of interference it creates for the successful recall of other items in the sequence. So one might suppose the perceived novelty (or repetition which reduces novelty) of intervening items may affect the profile of behavioural outcomes within the task.

Another, contrasting contemporary account of time-based memory encoding and retrieval is the idea that 'time for consolidation' enhances short-term memory (Bayliss, Bogdanovs, & Jarrold, 2015; Jolicœur & Dell'Acqua, 1998). During an unfilled time interval, memory consolidation can be defined as an active process that works to strengthen a new memory trace so that it can be successfully retrieved at a later point in time (Dewar, Alber, Cowan & Della Salla, 2014; Mercer, 2015). Arguably, visual memoranda suffer from rapid time based decay if there is reduced opportunity for engaging in a consolidation process (Knöchel et al., 2015); as a result, the memory trace is more vulnerable to interference or overwriting from succeeding items (Nieuwenstein & Wyble, 2014, offer a recent test of so-called short-term consolidation). One influential proponent of memory consolidation has been Wixted (2004). As this author points out, the concept is hardly new and may be traced to a forgetting law of Jost at the end of the 19th century captured by the insight that, with elapsing time, old encoded items within memory become less vulnerable to the disruptive effects of subsequent events: in other words they show an ever-slowing proportional memory decay function. This temporal gradient of retroactive interference, whereby allowing a temporal interval free of interfering material post encoding strengthens the memory trace, has intuitive appeal and empirical support. Thus, using visual characters (unfamiliar written items) as memoranda and varying time between items, Ricker & Cowan (2014) observed that limiting post-encoding consolidation time using brief inter-stimulus intervals, impaired memory recall. They concluded that "whether or not time-based forgetting will be observed in a working memory task is largely determined by the amount of time allowed for consolidation of working memory" (p. 427). Similarly, Bayliss et al. (2015) varied post-encoding time for lists of consonants by introducing a demanding processing activity either immediately or following a delay (whilst equating retention interval across conditions); they interpreted the impaired performance in the immediate condition as consistent with a consolidation process. Unfortunately, as a recent review (Ricker, 2015) of consolidation in short-term memory makes clear, there is a surprising lack of clarity as to the time-course of short-term consolidation or indeed whether or how it might differ from the more familiar 'encoding time' of the memory trace. It is, however, a factor that promises to be highly salient in any memory task using successive items separated by a certain inter-item interval – as in the BPS task.

1.8 The question of stimulus similarity

In human discrimination tests, stimulus ‘inputs’ may be varied by manipulating mnemonic similarity between item pairs so as to achieve an understanding of how sensitive observers are to discriminating small changes between items. It has been reported in the BPS – and this may not to be too surprising - that the proportion of correct ‘similar’ responses is lower for similar item pairs which have been rated high in perceptual similarity in separate tests (Kirwan & Stark, 2007). Notably, by varying the degree of item similarity Lacy et al. (2011) created a parametric scale of mnemonic similarity between object item pairs. This scale was based on the proportion of ‘old’ versus ‘similar’ responses to each item pair. Similar pairs frequently identified as ‘similar’ were regarded as reflecting large changes in input (low mnemonic similarity). Small changes in input resulted in item pairs incorrectly identified as ‘old’ (high mnemonic similarity). This scale of mnemonic similarity created by Lacy et al. has usefully been incorporated into the BPS and BPS-O task, with behavioural scores of pattern separation decreasing with increased mnemonic similarity (Stark, et al., 2015). The picture from the neuroscientific literature is consistent too (see General Discussion). Again, varying stimuli along a psychophysical continuum will be an important contribution of the present thesis.

It is worth bearing in mind that, whilst placing demands on pattern separation on the one hand and pattern completion on the other, the BPS measure has been criticised (Hunsaker & Kesner, 2013; Liu et al., 2016; Molitor et al., 2014) as it is based upon both encoding and retrieval processes and does not offer an independent (or ‘process pure’) window onto either encoding process. Consider that, when presented with highly similar items, a recognition memory task may bias participants towards a recall to reject strategy mentioned earlier in Introduction (this idea was noted by Gallo, 2004). Such circular reasoning clearly presents considerable difficulties of interpretation. Therefore, some suggest that removal of an explicit memory response, in favour instead of incidental encoding of similar items may be a more viable method of indexing pattern separation (Bakker et al., 2008; Lacy et al., 2011; Motley & Kirwan, 2012). Further, in a subsequent recognition phase, if we were to drop the similar response in favour of a two-response task (old, new) we might ‘get a handle’ on the same information (what level of fidelity is achieved during encoding?) by introducing a new experimental manipulation – namely varying the physical similarity across our set of items. In this way, errors of reporting old for similar stimulus items (varied across a continuum of similarity in the test items) would be all the information we might need to index the pattern separation mechanism.

1.9 The question of stimulus novelty

There is an impressive tradition of research in the neurosciences, represented for decades by the work of the great Russian neuroscientist Olga Vinogradova (a 2011 paper prints some historically interesting material) and others where the hippocampus has been central to models of human and animal learning and memory – and identifying the novel in the environment (indeed Vinogradova referred to the hippocampus as a novelty detector). Nevertheless specifically behavioural evidence of pattern separation and completion processes in humans is still in its infancy (for a recent review, see Deuker et al., 2014; also Yassa & Stark, 2011). Indeed, in a recent systematic review of behavioural studies of pattern separation (Liu et al., 2016) only sixty-two studies were identified as providing evidence of pattern separation and pattern completion processes in sensory/perceptual memory, with just a handful of these satisfying certain reasonable criteria of acceptability (this point will be elaborated upon below).

1.10 Tasks, measures and analyses

The raw responses in the three-response task are straightforward, but their interpretation – and most especially the interpretation of the derived measures (such BPS and BPC) – are far from straightforward. The approach adopted in this thesis is to cautiously adopt the most standard derived scores, remind the reader throughout that their validity is in question, and then in General Discussion attempt a resolution or at least confront head-on the validity of the derived scores. Intuitively – what may be termed the face validity of the task and its measures – the three-response is tapping discrimination and generalisation, the two fundamental mental operations underlying the separate encoding of neighbours in feature space into memory, the reliable retrieval of the most appropriate representation, and the identification of novelty in the environment. There are a number of problems, however. The first difficulty is that performance patterns across accurate and inaccurate new, old and similar response types, captured by the derived measures do not permit independence across these measures; a second difficulty is deciding on the most appropriate false alarm measure (and this may simply not be resolvable); a third and arguably more challenging difficulty is the distribution of response bias across a three-choice task.

In the standard two-choice test, the decision axis can be placed between the two choices, and the calculation of a bias to one can be made. Unfortunately, in BPS the partition of response bias or criterion does not fall comfortably within one decision space (in signal detection theory, under the receiver operating curve). Rather, consideration of each measure invokes not a single criterion of responding, but two criteria (for example, a bias to respond ‘old’ and a bias to report ‘similar’). Clearly these are relatively deep waters statistically, and the common assumption within signal detection when there are more than two alternatives is to side-step response bias – assume there is no bias in responding. Here it may or may not be reasonable to assume that, prior to testing, our participants have no bias to report one thing rather than another, and indeed in the literature ignoring bias for the somewhat similar m-AFC cases (three stimuli presented but with one response) has been common because of the challenges of interpretation (Wickens, 2002; but see DeCarlo, 2012, discussed in General Discussion Section 8.7). Luce (1963) stated ‘‘The generalization of the two-alternative signal detectability model to the k-alternative forced-choice design is comparatively complicated if response biases are included and very simple if they are not’’ (p.137). Yet, as Macmillan and Creelman note (2005, p. 250), the fact that bias is customarily ignored for multiple-choice (> 2) designs does not mean it disappears. They are referring to the m_AFC arrangement, and of course the continuous recognition situation only offers a single item on each trial, but the principle of the three choices is very similar. Still, the present thesis has taken that course and used z-transformed BPS and BPC derived scores similar to d’ throughout as a best compromise.

Therefore, three outcome measures were calculated: the recognition accuracy score (RAS), the behavioural pattern separation (BPS), and the behavioural pattern completion (BPC) using signal detection theory transforms (Z). That is, these derived measures were based on hits and false alarms across the stimulus and response possibilities in the task (shown in Table 1.2) following Stark et al. (2013), but using signal detection methods; our measures used a normalized score as a bias free index of sensitivity (d', Bi, 2002). The RAS equalled the mean normalized proportion of old items correctly identified as 'old' minus the mean normalized proportion of new items incorrectly identified as 'old'. The BPS equalled the mean normalized proportion of similar items identified as 'similar' minus the mean normalized proportion of new items incorrectly identified as 'similar'. The BPC equalled the mean normalized proportion of similar items incorrectly identified as 'old' minus the mean normalized proportion of new items incorrectly identified as 'old'. Hits and false alarm rates for each derived measure were corrected for floor and ceiling effects (Macmillian & Creelman, 2005; Macmillian & Kaplan, 1985)[1.] Throughout the experiments reported in this thesis analysis is based on these derived measures. In addition to this, just occasionally the raw proportions uncorrected for response bias are selectively reported to allow for the full examination of changes to the derived measures. The assumptions underlying the measures will be revisited in General Discussion, but the reader will be reminded of the difficulties of comparison of measures throughout. The important point here, is that it is believed that the measures are capturing the essential properties of memory sensitivity (for example, retrieving a trace from a noisy representational background), and the other side of the signal detection coin *response bias*. As explained by Stark et al. (2013) the key derived BPS score is explicitly designed to take account of response bias, since they assess pattern separation performance by calculating the ratio of similar responses given to similar lures minus similar responses given to novel

---

[1] Extreme cases (hit rate of 1 and false alarm rate of 0) were adjusted using the correction 1-1 (2N) and 1/(2N) with N as the number of trials on which the proportion in based on (Macmillian & Creelman, 2005).

foils to account for any similar response bias overall; an observer having poor pattern separation performance will show low BPS scores because they will make fewer similar responses to similar lure trials (typically making more old responses). In Chapter 6, with some assumptions cautiously if not fully confidently made, a criterion bias measure C is calculated for participant data in a two-response version of BPS (See Chapter 5, Section 6.5.2). In General Discussion (Section 8.7) the question of criterion bias will be revisited.

Table 1-2.

Definitions of the response possibilities for the new, old and similar stimuli; and formulae used for the main derived measures.

	Response		
Stimulus	New	Old	Similar
New	Correct rejection Rate	False alarm rate	Similar bias rate
Old	Miss rate	Hit rate	Incorrect
Similar	Incorrect	Pattern completion rate	Pattern separation rate

Recognition Accuracy Score (RAS) = Z(hit rate) – Z(false alarm rate)

Behavioural pattern separation (BPS) = Z(pattern separation rate) – Z(similar bias rate)

Behavioural pattern completion (BPC) = Z(pattern completion rate) – Z(false alarm rate)

Chapter 2 The question of clinical measures of cognitive impairment

The BPS is already ‘out there’ – it is currently adopted within clinical settings, as a neuropsychological test used to identify the early behavioural markers of cognitive impairment related to neuropathological conditions, notably amnesic mild cognitive impairment (aMCI; Yassa et al., 2010) and Alzheimer’s disease (AD; Stark et al., 2013). It is claimed that BPD is able to differentiate between aMCI, AD and normal ageing. To achieve this, as outlined in Introduction, the critical study manipulation has been to vary the number of intervening items between 'old' and 'similar' item (and ‘old’ and ‘new’ item) presentations (Ally et al., 2013). Ally et al. report pattern separation performance in healthy older adults, aMCI and AD groups, at differing numbers of intervening items (4, 12 or 40), a variable the authors termed 'lag'. As expected, pattern separation performance was lowest for AD patients. However, in comparison to healthy older adults, pattern separation performance in the aMCI group decreased with increasing lag, whilst AD showed flat, poor or chance performance at each lag. Ally et al. make the claim that performance in aMCI is evidence of rapidly degrading visual representations in immediate memory. Unfortunately the lag manipulation introduces confounds which are examined below.

2.1 Experiment 1

The present experiment tests the idea that the lag manipulation by Ally et al. (2013) introduced two confounds, namely ‘time since encoding’ (the decay interval) and ‘post-item delay’ (an interval permitting time for consolidation of items), in their manipulation of number of distractors. Here participants completed the continuous recognition task using computer-presented pictures of everyday objects as used by these authors (available from the BPS on-line resources). Since the interest is in the validity of the lag manipulation (whatever the population studied), a younger “cognitively unimpaired” sample was studied. Stimulus pairs ('old' - 'old' and 'old' - 'similar' pairings) were separated by either 4 or 12 intervening items. In order to separately examine any effects of the retention interval itself, the memory retention interval between these stimulus pairs was manipulated. This necessitated the introduction of a third critical factor within the memory task: the inter-item interval, which is the temporal interval between successively presented items within the continuous recognition task. Therefore, there were two temporal intervals for us to consider: firstly, the gap between each successive item; and secondly, the gap between a first occurrence of an item and the time of its repeated presentation later in the series of trials (for ‘old’ pairings), or the time of a similar pairing (for ‘similar’ pairings). It is this second, memory retention interval that is ‘confounded’ by the number of intervening items, and therefore manipulating this factor is a key test and a novel one. Critically it is arranged in two experimental conditions to maintain the exact same memory retention interval but introduce either 4 or 12 distractor items within that interval. One key prediction, arising out of consideration of contemporary accounts of forgetting in immediate memory, is that performance on the derived measures will benefit from extended inter-item spacing, allowing post-item time for consolidation into immediate memory; a second expectation is that intervening items may produce item-specific as well as item-nonspecific interference; finally, acknowledging time-based trace decay, the longer the retention interval the poorer the performance (one condition has a much shorter memory retention interval, whilst having the same number of distractors in that interval as one of the longer retention interval conditions). Since we are interested primarily in detection performance on the task generally, the key focus will be upon what may be termed the ’standard’ measure of d’ in the task, namely the recognition accuracy score. The other two measures, BPS and BPC are analysed for completeness but RAS is the key one.

2.2 Method

2.2.1 Participants

Thirty-five younger adults (32 female), aged between 18 - 22 years (M_{age} = 19.56 years, SD_{age} = .88) were recruited from the School of Psychology at the University of Leeds. Participants were all native English speakers, with self-reported normal or corrected-to-normal vision. Prior to recruitment, sample size was decided using G*power analysis (Faul, Erdfelder, Lang, & Buchner, 2007). For the present task, 28 participants were needed to find a medium effect size (f= .25) with α= .05, 1-β= .80 and a moderate correlation between repeated measures of r= .50. Prior to Experiment 1, half of the participants (n=17) completed experiment 2. There was no experiment order effect for the overall proportion of correct responses, $t(32) = .19$, $p = .849$. Ethical approval for Experiment 1-2 was granted by the University of Leeds ethics and research committee (Ref: 13-0289).

2.2.2 Materials

Visual stimuli were coloured photographs of common everyday objects. The everyday object stimuli were by permission of Dr. Craig E. Stark, University of California. From the set, 90 similar pairs and 270 individual items were used as 'new' and 'old' items. Each item was adjusted to a height of 227 pixels and width of 178 pixels. Visual stimuli were presented on a white background, on a Dell 1708FP monitor, with participants approximately 60cm away from the monitor. The experiment was run using E-Prime 2.0 software (Schneider, Eschman, & Zuccolotto, 2002).

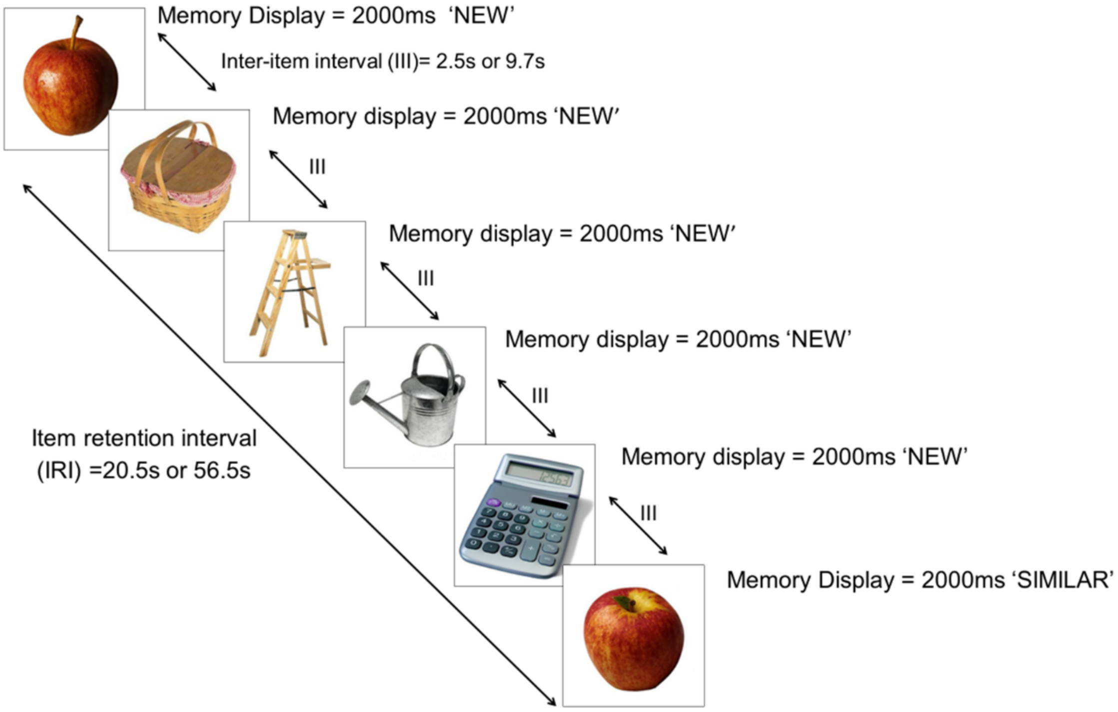

Figure 2-1. Example of the experimental procedure in the lag conditions: 4_short and 4_long.

2.2.3 Design and procedure

The continuous recognition task (adapted from Ally et al., 2013) was arranged into 60-trial blocks, consisting of 20 single 'new' items, 10 'old' item pairs (an item presented and later presented again), and 10 'similar' item pairs (an item presented and a very similar item presented later). Note the use of ‘new’, ‘old’ and ‘similar’ to refer both to our stimuli and to the responses of our participants. It is evident where a word refers to one or other, and so we avoid the confusion of the overabundance of terms (new, similar, lure, etc.) common within the literature. Yet it is evident when a word is referring to a memory item or a response. This decision was made in response to the overabundance of terms within the pattern separation literature, generating confusion when comparing key outcome measures.

As shown in Figure 2.1, each item was presented on the screen for 2 s, followed by an inter-item interval (III) of either 2.5 s or 9.7 s. Old and similar item pairs were separated by 4 or 12 intervening items. A repeated measures design was employed with the number of intervening items and III as independent variables. Participants completed 3 item lag conditions: 4_long (4 intervening items with a 9.7 s III), 12_short (12 intervening items with a 2.5 s III), and 4_short (4 intervening items with a 2.5 s III). Based on duration and III, the item pairs (old and similar) were separated by an item retention interval (IRI) of either 20.5 s (4_short) or 56.5 s (12_short and 4_long) intervening items). Of importance is the equal IRI between the two item lag conditions, 4_long and 12_short.

For each item presented in the current sequence of items participants were instructed to report by button press whether it was new (first time presented), old (the item had been previously presented) or similar (item shared similar but not identical visual features, to an item previously presented). Participants completed nine experimental blocks divided across two sessions, both in counterbalanced order. Each session began with 12 practice trials. Following this, participants completed three blocks for each condition (a total of 180 trials). There was a 10 minute break between blocks. Each item lag condition was completed in separate blocks, lasting either 11.7 min (4_long) or 4.5 min in duration (4_short and 12_short). The number of experimental block completed in each test session was adjusted according to item lag condition, in order to maintain approximately equal duration. The order of condition was counterbalanced across the sessions. Visual items were not repeated across experimental blocks and this was emphasised in the instructions provided to participants. They were encouraged to determined their judgments of old, new and similar on the current sequence presented.

2.3 Results

2.3.1 Overview of measures

The mean proportion of responses was calculated for lag (4_long, 12_short, 4_short). For each stimulus condition, the mean proportion of each possible response (new, old and similar) is shown in Table 2.1. Outliers were identified based on the proportion of new responses for old and similar items[2]. One participant was excluded as their 'new' responses exceeded 60 % new responses for old and similar items (*N* = 34). Note our main interest is standard discriminative performance, our RAS measure, but for completeness some subsidiary analyses are briefly reported. To attempt to control for Type 1 errors in the *t*-test comparisons, the more conservative *p* value of .01 was adopted (although some marginal p values are also reported).

---

[2] Across the three item lag conditions, participants' ability to accurately identify new items was high. A repeated measures ANOVA with a factor of lag (4_long, 12_short and 4_short) revealed no main effect of lag for the correct rejection rate, $F(2, 66) = .25$, $p = .782$, $\eta_p^2 = .01$.

Table 2-1.

Proportions (p) for each response possibility (New, Old, and Similar) across lag for object stimuli. Standard deviation is given in parentheses.

Stimulus		Lag 4_long	 12_short	 4_short
New	Correct rejection rate	.96 (.04)	.96 (.06)	.96(.06)
	False alarm rate	.02 (.03)	.02 (.03)	.02 (.04)
	Similar bias rate	.02 (.03)	.02 (.04)	.02 (.03)
Old	Miss rate	.07 (.09)	.11 (.10)	.19 (.26)
	Hit rate	.85 (.11)	.81 (.11)	.74 (.26)
	Incorrect	.08 (.08)	.08 (.07)	.07 (.08)
Similar	Incorrect	.11 (.12)	.15 (.14)	.16 (.13)
	Pattern completion rate	.41 (.16)	.43 (.17)	.38 (.16)
	Pattern separation rate	.49 (.20)	.43 (.19)	.46 (.19)

2.3.2 Performance on the task as RAS score

The outcomes for the key RAS measure are shown in Table 2.2 (for completeness BPS and BPC are also shown). A one way repeated ANOVA with a factor of lag (4_long, 12_short and 4_short) was not significant, F (2, 66) = 2.60, p = .082, η_p^2 = .08. However, pairwise comparisons revealed RAS was marginally greater for 4_long than 4_short, t (33) = 2.15, p = .039, d = .45; 12_short and 4_long did not differ, t (33) = 1.35, p = .185 ; 12_short and and 4_short did not differ, t (33) = 1.17, p = .252.

2.3.3 Performance on the task as BPS and BPC scores

There was no main effect of lag for BPS, *F (*2, 66) = .41, *p* = .67, η_p^2 = .01. For BPC, lag was significant, *F (*2,66) = 3.38, *p* = .040, η_p^2 = .09. Paired t test revealed a non-significant difference for BPC between 4_long than 4_short, *t* (33) = 1.87, *p* = .071, *d* = .31; although a marginally higher BPS score for 12_short than 4_short, *t* (33) = 2.37, *p* = .024, *d* = .45.

Table 2-2.

Mean recognition accuracy score, behavioural pattern separation score and behavioural pattern completion score as a function of lag. Standard deviation given in parentheses.

	Lag		
	4_long	12_short	4_short
Recognition accuracy score	3.28 (.71)	3.12 (.56)	2.92 (.87)
Behavioural pattern separation score	2.08 (.69)	1.99 (.74)	2.05 (.75)
Behavioural pattern completion score	1.88 (.48)	1.95 (.50)	1.74 (.45)

2.4 Discussion

This opening experiment focused on a simple discriminative measure of performance on the three-response task, namely a standard d’ sensitivity measure, calculated as the z score transformed ‘hit rate’ (reporting ‘old’ items correctly as ‘old’) minus the z score transformed ‘false alarm rate’ (reporting ‘new’ items as ‘old’). This is referred to as Recognition Accuracy Score (RAS). This, it is argued, is the least ambiguous of the various proposed transformed or derived indices of performance in the BPS task, having a single response-bias dimension. It is an attractive measure for present purposes – the interest is in how temporal and inter-item factors influence basic discriminative performance or sensitivity in the task. A summary (if tentative, given the marginal statistical outcomes) conclusion from this first experiment is that participants’ responses to both old and similar items was not driven by the level of interference between encoding and recalling an item pair, but may be governed by the post-item interval. Within a fixed retention interval, performance in our RAS measure did not significantly decline as the number of intervening items increased from 4 to 12. However, when the number of intervening items was equal (at 4 intervening items), an increasing inter-item interval (from 2.5 to 9.7 seconds) improved discriminative performance. No claim is made that proactive or build-up interference plays no role in the BPS (proactive interference is well documented in situations where several stimuli are presented in sequence in these sorts of task; (c.f. Makovski & Jiang, 2008; but see Oberauer, Awh, & Sutterer 2017), but at least across the stimulus parameters and number of items used here, the more important variable appears to be the spacing of items. This spacing may permit enhanced opportunity for encoding through a form of short-term memory consolidation.

Now consider the time manipulation. Two conditions were arranged to have differing numbers of intervening items (4 or 12) but identical and quite extended memory retention intervals of 56.5 seconds (4_long and 12_short). Decay of the item memory trace in immediate memory, although unpopular with many contemporary theoreticians (e.g. Farrell et al., 2016), is the standard description of forgetting occurring when a retention interval is extended (Ricker et al., 2014). If we wish to think only of the memory trace retention interval, the critical conditions in Experiment 1 are 4_short, where it had a duration of 20.5 seconds, and 4_long, here it had a duration of 56.5 seconds. By the three measures there was no indication of trace decay; indeed if anything, marginally higher performance on RAS was noted in the 4_long than the 4_short condition. In addition, the BPS score was marginally higher for 12_short than 4_short, that is better performance in the much more extended retention interval condition.

However, it may be acknowledged that the present manipulation of intervening items and inter-item delay failed to perfectly experimentally disentangle these variables. The difficulty was that an extended 9.7 second delay in a 12 intervening item condition would have led to unacceptable stimulus session durations. Keeping this constraint in mind, however, it is still the case that ‘time for consolidation’ enhanced discriminative performance despite a longer item retention interval. Therefore, the present test of short-term consolidation time was impressive – it shows itself here in discriminative performance despite a greater opportunity for temporal trace decay. However, the further discussion of how time and number of items may play their role in BPS will be postponed to the next chapter, which offers arguably a more valid window onto performance changes when these variables are manipulated – for in the next experiment a worrying factor is addressed, namely the opportunity of the participants to engage in simple verbal rehearsal of the memoranda. For now, we may echo a statement in a recent review paper (Yeal & Yassa, 2018) that it is probably premature at this stage to apply the three-response behavioural task in making clinical distinctions, such as susceptibility to effects of numbers of distractor items.

Chapter 3 The question of developing valid stimuli

3.1 Experiment 2

A clear feature in the first experiment was that, across all three item lag conditions performance, based on both the RAS and BPS scores, was surprisingly high overall. As explained in Section 1.5 of Introduction, this high level might be attributed to the availability of verbal labelling of the stimulus items (such as ‘cup’, ‘apple’, ‘chair’). In recent published reports, the use of familiar or everyday visual stimuli has encouraged verbal encoding of stimuli as well as the activation of long term representations from memory (Liu et al., 2016; Mckeown et al., 2014). This is clearly a worrying confound - if the participant engages in sub-vocal rehearsal processes to maintain the memory trace we are not observing stimulus representations’ encoding and retrieval per se but rather their translations into a verbal code (and a code that may be an activated long-term code formed some time before the participant enters the experimental room). Indeed, it is known that, for mixed-category visual stimuli, verbal labels can greatly improve the retention of items across extended time intervals (Lupyan, 2008; Richler, Palmeri, & Gauthier, 2013); the usual account stresses that verbal encoding or translation counteracts decay and interference through active rehearsal (Berman, Jonides, & Lewis, 2009). Note also, in Experiment 1 the 4_long condition greatly extended the retention interval, but performance appeared insensitive to this – in other words no temporal decay was evident; verbal encoding and maintenance countering such putative decay might explain this.

So the use of verbal labels can be an effective strategy to improve performance in visual recognition memory experiments, as categorising memoranda forms a less noisy representation (Lupyan, 2008) – but it does not really tell the experimenter much about visual recognition memory per se. Of course, experimenters have long known that it might be advantageous to limit or prevent such verbal encoding or rehearsal strategies in their participants; and the most common method of preventing active rehearsal of memory items has been to introduce a secondary task (Lewandowsky, Geiger, Morrell & Oberauer, 2010). For instance, verbal rehearsal is often reduced or even eliminated by introducing articulatory suppression during the memory retention interval (Lewandowsky & Oberauer, 2015). Unfortunately, in doing so, the experimenter may be inadvertently introducing additional interference during the retention interval – the secondary task itself may be retroactively interfering with the maintenance of the memoranda (Dewar et al., 2014). One solution may be to side-step the problem by adopting stimuli which do not lend themselves to verbal encoding or translation. For example, it would be difficult to quickly generate and rehearse a verbal label for a random ‘squiggle’ on a piece of paper. A second solution (and this does not preclude the first) is to use a set of stimuli that broadly fall into the same verbal category. Therefore, if all the items in the study session are squiggles, it will clearly not be of much help to rehearse the word squiggle. This second method of controlling rehearsal was the one adopted here. A new set of visual stimuli was used with a single label – doors. In Experiment 2, the effect of intervening item and inter-item interval variables were again examined using the same continuous recognition task arrangement as in Experiment 1, but with the memoranda being a large set of photographs of doors. The aim of this experiment is foremost to design and evaluate what is arguably a more valid stimulus set for the BPS and a key interest is how overall performance will respond to the removal of verbal encoding or maintenance strategies (hopefully pattern separation and pattern completion processes will be revealed more purely). However, the experiment also offers another chance to test the claims of Ally et al. (2013), so the same manipulations are made here of number of intervening items and inter-item interval. Recall, in Experiment 1 there was modest evidence that an extended period of time between items enhanced discrimination (RAS) performance. One cautious statement in discussion of that experiment was that the stimuli had permitted verbal labelling and verbal maintenance, hiding therefore the sort of time sensitive effects that were of central interest (such as decay of the memory trace). So the new stimuli will provide a more realistic test of the time-based encoding idea on the one hand (opportunity for post-item short-term consolidation) and the time-based decay idea on the other (note the 4_long and 12-short conditions in Experiment 1 had the most extended memory retention interval). Since, however, the bigger question is improving the validity of the BPS, the core literature derived measures, BPS and BPC, will be of at least equal interest to the RAS. The predictions for the manipulations mirror those made in the first experiment, namely that an extended retention interval should manifest as declining performance (time based decay), but this may be offset by an extended inter-item interval (reduced proactive interference and increased opportunity for post-item short-term consolidation). Whilst Experiment 1 failed to demonstrate time-based decay, it was argued that this may have been offset by the opportunity for verbal labelling and maintenance; the new stimulus set should counter this. So we predict highest discriminative performance (by the RAS measure) in 4_long, intermediate performance in 4_short, and lowest performance in 12_short (the latter having both the longest memory retention interval and the greatest opportunity for inter-item interference). Regarding the BPS and BPC measures, the key interest is in performance generally – the question is, when we remove the opportunity for verbal labelling, does a younger cognitively normal sample demonstrate high levels of pattern separation and pattern completion? Otherwise stated, do the new stimuli offer a promising improvement on the everyday items common in the literature? Naturally, it will also be of interest to explore how the BPS and BPC performance outcomes mirror the RAS measure (bearing in mind that these three derived measures are not directly comparable, or indeed, independent assays of memory performance on the task – see below).

3.2 Method

3.2.1 Participants

Thirty-five younger adults (32 female, M_{age} = 19.56 years, SD_{age} = .88), native English speakers, had self-reported normal or corrected-to-normal vision. Prior to the experiment, half of the participants (*n* = 18) completed Experiment 1.

3.2.2 Materials

The Materials were the same as in Experiment 1, except for the following changes. Stimuli were coloured photographs of a wide variety of doors (door stimuli). The stimuli were taken from a larger item database3, provided by Professor Alan Baddeley, University of York. From the set, 90 similar pairs and 270 individual items were used as 'new' and 'old' items. Each item was adjusted to a height of 227 pixels and width of 178 pixels. Lacy et al. (2011) outline in detail the degree of similarity between the object stimuli used in the BPS task. Traditionally, the similar item pairs are two different objects that are matched in terms of their category label and perceptual similarity for example, two pictures of hammers that share similar visual features. In line with this, the new door similar item pairs were matched by Likert rating scale by a separate group of participants (*N* = 29, see below) according to category and perceptual similarity. The door stimuli were selected from a larger database of 400 pictures. This database consists of 100 groups of 4 items that are matched according to the following category features: function, age, colour, glazing, condition, shape, opening, details,

---

[3] The database has since been published in Baddeley, Hitch, Quinlan, Bowes, & Stone (2016).

surround and memorability. To obtain a set of similar item pairs that matched (ideally) in both category features and perceptual features, two independent researchers generated every possible pairing within each door category group to select 600 item pairs. These item pairs were then rated for perceptual similarity using a 7-point Likert scale (1 = weak similarity and 7 = strong similarity). Based on the mode similarity rating, 90 similar item pairs were selected to be used in the main experiment.

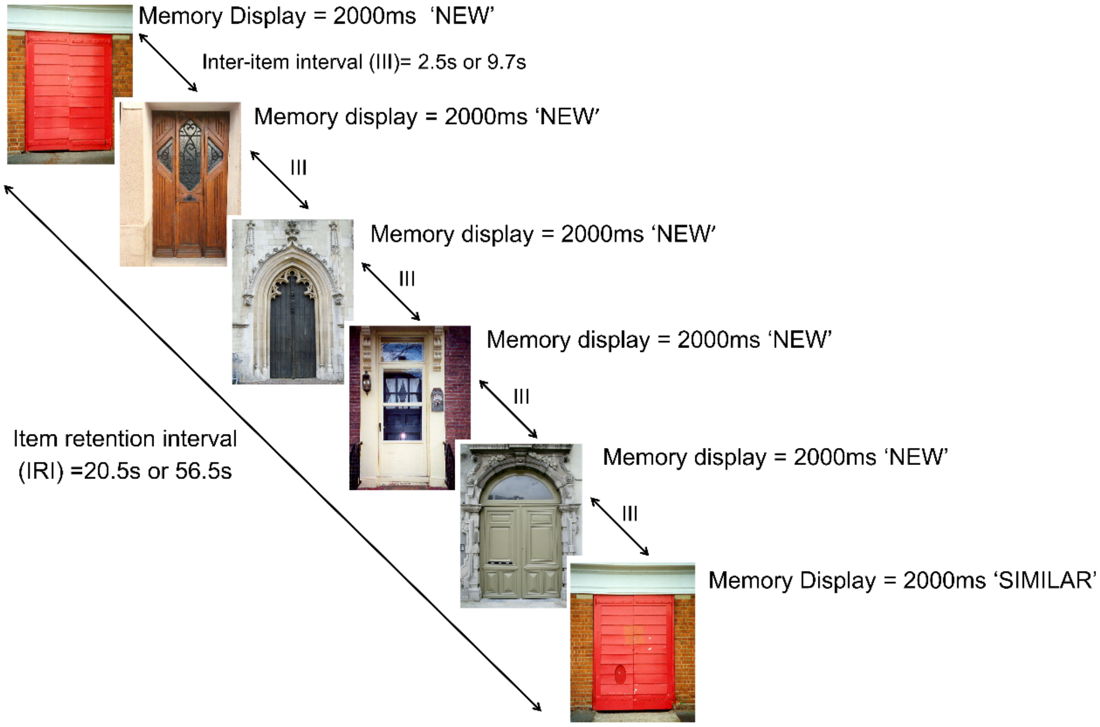

Figure 3-1. Example of the experimental procedure in the lag conditions: 4_short and 4_long for experiment 2.

3.2.3 Design and procedure

Participants completed a version of the continuous recognition task with the door stimuli. The design of the continuous recognition task mirrored Experiment 1. As shown in Figure 3.1, the number of intervening items separating old and similar item pairs, as well as the inter-item interval between each item, was once again manipulated. This produced three item lag conditions; 4_long (4 intervening items with a 9.7 s III), 12_short (12 intervening items with a 2.5 s III), and 4_short (4 intervening items with a 2.5 s III). Participants completed two sessions with the order counterbalanced, as was the order of condition within a session.

Consistent with previous research (Lacy et al., 2011) a measure of perceptual similarity for the new door stimuli was first calculated for each similar pair. Twenty nine younger adults (M_{age} =18.97, SD_{age} = 2.04), who did not participate in the continuous recognition task, viewed 90 selected door item pairs, and were asked to rate their perceptual similarity on a 7-point Likert scale (1 = weak similarity to 7 = strong similarity). The ratings revealed no significant difference in perceptual similarity between pairs used across our experimental conditions (χ^2 (12, N = 90) = 9.5, p = .66).

Table 3-1.

This reproduces Table 1.2 in Introduction, and shows definitions of the response possibilities for the new, old and similar stimuli; and formulae used for the main derived measures.

Stimulus	Response		
	New	Old	Similar
New	Correct rejection Rate	False alarm rate	Similar bias rate
Old	Miss rate	Hit rate	Incorrect
Similar	Incorrect	Pattern completion rate	Pattern separation rate

Recognition Accuracy Score (RAS) = Z(hit rate) – Z(false alarm rate)

Behavioural pattern separation (BPS) = Z(pattern separation rate) – Z(similar bias rate)

Behavioural pattern completion (BPC) = Z(pattern completion rate) – Z(false alarm rate)

3.3 Results

3.3.1 Overview of measures

The mean proportion of responses was calculated for each item lag (4_long, 12_short, 4_short), shown in Table 3.2 (*N* = 34: one participant identified as an outlier was excluded as their Miss rate/Incorrect exceeded 60 % of new responses) in order to derive the RAS, BPS and BPC (Table 1.2 from Introduction is reproduced below as Table 3.1 should the reader wish to be reminded of how the derived scores are arrived at). Analyses below are based on these derived measures. An important point worth reiterating here, is that whilst RAS conforms to the familiar form for a two-choice signal detection calculation (based on unambiguous hits and false alarms), BPS and BPC do not. It may be desirable to base our derived scores on z normalised scores, and follow the usual proposal for the choice of score to arrive at the measure, but please bear in mind that the derived measure calculations for BPS and BPC violate the basic assumption of signal detection theory of the reciprocal relation between ‘hit’ and ‘false alarm’ (see General Discussion for further elaboration on this constraint); casually, the transforms for RAS and the other two measures cannot validly be said to be the same sort of thing so we label RAS as a *d*’ measure, and the others as z transformed measures. Further statistical tests were performed on the raw scores to see whether they revealed something of interest, and this is noted below.

Table 3-2.

Mean proportions (p) for each response possibility (New, Old, and Similar) across lag for door stimuli. Standard deviation is given in parentheses.

		Lag		
Stimulus		4_long	12_short	4_short
New	Correct rejection rate	.82 (.12)	.78 (.14)	.79 (.16)
	False alarm rate	.07 (.06)	.09 (.08)	.07 (.05)
	Similar bias rate	.12 (.10)	.13 (.11)	.14 (.13)
Old	Miss rate	.13 (.13)	.24 (.19)	.16 (.18)
	Hit rate	.79 (.17)	.64 (.22)	.71 (.21)
	Incorrect	.08 (.08)	.12 (.10)	.09 (.11)
Similar	Incorrect	.29 (.18)	.39 (.20)	.41 (.21)
	Pattern completion rate	.44 (.15)	.39 (.19)	.34 (.16)
	Pattern separation rate	.27 (.14)	.22 (.13)	.25 (.15)

3.3.2 The RAS

For RAS, a one way repeated ANOVA with factor of lag was significant *F (*2, 66) = 19.2, *p* < .001, η_p^2 = .37. Paired sample t-tests showed performance on this measure was higher at 4_long (*M* = 2.64, *SD* = .78) than for 12_short (*M* = 1.93, *SD* = .74), *t* (33) = 5.55, *p* < .001, d = .94; 4_long was also higher than 4_short (*M* = 2.28, *SD* = .64); *t* (33) = 3.57, *p* = .001, *d* = .52. RAS for 4_short was marginally (by the adjusted p value for multiple t-tests) higher than for 12_short, *t* (33) = 3.08, *p* = .004, *d* = .51. Thus, the outcomes for RAS followed predicted patterns for benefits of post-item interval, but a mixed or uncertain outcome for memory retention interval (recall that the 4_long and 12_short conditions had the more extended 56.5 second retention interval).

3.3.3 The BPS

For BPS a one-way repeated measures ANOVA with factor of lag was significant, *F (*2, 62) = 8.51, $p < .001$, $\eta_p^2 = .22$. Paired sample t-tests showed performance on this measure was higher for 4_long (*M* = .76, *SD* = .55) than for 12-short (*M* = .46, *SD* = .41); *t* (32) = 3.34, *p* = .002, *d* = .64; and marginally higher than 4_short (*M* = .6, *SD* = .49); *t* (31) = 2.15, *p* = .04, *d* = .31; also, 4_short was marginally higher than 12_short, *t* (32) = 2.21, *p* = .034, *d* = -.32. The performance patterns therefore follow the RAS measure. However, a notable feature of the scores are that they are low.

3.3.4 The BPC

For BPC a one-way repeated measures ANOVA with factor of lag was significant, *F (*2, 66) = 9.23, $p < .001$, $\eta_p^2 = .22$. Paired sample t-tests showed performance on this measure was higher for 4_long (*M* = 1.48, *SD* = .34) than for 12_short (*M* = 1.21, *SD* = .5); *t* (33) = 2.92, *p* = .006, *d* = .65; and 4_short (*M* = 1.15, *SD* = .41); *t* (33) = 5, $p < .001$, *d* = .89. No other pairwise comparison was significant. The performance patterns are therefore again in line with the RAS, supporting benefits of post-item interval. The scores overall are, like the BPS, somewhat modest

Table 3-3.

Mean Recognition Accuracy Score, Behavioural Pattern Separation score and Behavioural Pattern Completion Score, for each lag (4_long, 12_short, 4_short)

	4_long	12_short	4_short
Recognition Accuracy Score (RAS)	2.64 (.78)	1.93 (.74)	2.28 (.64)
Behaviour Pattern Separation Score (BPS)	.76 (.55)	.46 (.41)	.60 (.49)
Behaviour Pattern Completion Score (BPC)	1.48 (.34)	1.21 (.50)	1.15 (.41)

3.4 Discussion

Using a continuous recognition task with a large stimulus set of single-category memoranda, the present experiment aimed to examine separable effects of inter-item interval, number of intervening items and item retention interval. By manipulating the inter-item interval, it was possible to maintain a fixed item retention interval between old and similar pairs, whilst independently manipulating the number of items intervening within that interval. Participants' ability to identify similar and old items was strongly affected by the manipulation of the post-item interval. Importantly, the BPS (thought to reflect a form of pattern separation in immediate memory), the BPC (thought to reflect a form of pattern completion in immediate memory) and RAS (thought to reflect broadly the fidelity of the memory trace or sensitivity in the task) consistently showed declining performance as the post-item interval was reduced from

9.7s to 2.5s. In the 4_short and 4_long conditions, where the number of intervening items was equal, performance improved with increasing post-item intervals (and this despite a very much extended memory retention interval in the 4_long condition). The explanation offered here (and pursued in General Discussion) is that the longer duration post-item interval promotes short-term memory trace consolidation (an initial encoding benefit). In contrast no appeal to forgetting through interference (number of distractors in the retention interval) or through memory trace decay (extending the memory retention interval) may be made in describing the present data.

Across the first two experiments, three standardised measures of recognition accuracy, behavioural pattern separation and behavioural pattern completion have been reported. In Experiment 2, it is noteworthy that the proportion of 'similar' responses to similar items, thought to index pattern separation and termed pattern separation rate (therefore, presumably precise separable representations of similar items within memory) is quite low (quite commonly reported elsewhere, for example Toner et al., 2009). It may be difficult to see how the presumed underlying processes 'trade' against one another. That is, across the three-response profile in the BPS task there may be some trading between proportions correct (or conversely errors) on one index and proportions correct or errors on another. For example, it is perfectly conceivable that 'errors' in the form of reporting similar target items as old (a failure of pattern separation) would be 'balanced' by correct reports of old items as old (a demonstration of pattern completion). This latter raw score, termed hit rate, is indeed high in both Experiment 1 and Experiment 2. Arguably the task inherently confounds attempts to perfectly separate the putative encoding and retrieval stages. Authors have indeed identified this concern (e.g. Liu et al., 2016; Molitor et al., 2014).

Another notable feature in the data is that performance for similar items is low; indeed accurately identifying similar items (termed pattern separation rate) is low in both these first experiments (Tables 2.1 and 3.2). At the present, there is a lack of clarity in the pattern separation literature in regards to the process by which participants are able to identify similar items. One suggestion is that participants engage in a recall to reject strategy (Kirwan & Stark, 2007). In study/test recognition paradigms, a recall to reject strategy might be employed when the target and non-target (new) items share categorical or perceptual similarities (Gallo, 2004). Casually, in memory tests for series of pictures, to ensure accurate responses to similar items, participants may attempt to recall the stored memory of the initial item (Kirwan & Stark, 2007; Morcom, 2015) and then compare the degree of similarity between the stored item and any similar items presented later. If the observer fails to identify the mismatch (Duncan et al., 2012) or the degree of similarity is too high (Lacy et al., 2011), pattern completion processes are engaged and similar items are incorrectly identified as old. This notion will be revisited in the following chapters.

Chapter 4 The question of ageing

This chapter extends the investigations in the preceding chapters with a sample of healthy older adults. This is inspired by reports using the continuous recognition task, that older adults show impaired object recognition in comparison to younger adults (Yassa et al., 2011). Specifically, for the three-response task, it has been shown that, whilst performance between younger and older adults is not different for old and new items, older adults demonstrate a reduced ability to accurately identify similar items (Toner et al., 2009). Notably, older adults show more incidences of false recognition by incorrectly identifying similar items as ‘old’ (Stark et al., 2013). The results reported in this chapter further support the view that extending post-item interval is beneficial, particularly for healthy older adults – at least for some measures. Results point to an age difference in efficiency during encoding.

4.1 Experiment 3

The present experiment was guided by the considerations outlined in Introduction Section 1.6 on how recognition memory and particularly sensitivity to small changes in ongoing stimulation, may decline with age. Also revisited is the question of the interplay between ‘time for consolidation’ and ‘time to recall’ explored in the first two experiments. The same three-response task was employed as in the first two experiments, again using the door stimuli. Participants were healthy older adults. Within the continuous sequence of items (pictures of doors), old and similar item pairs were separated by either 4 or 12 intervening items. Again, by manipulating the inter-item interval between successive items, it was possible to maintain an equal retention interval between item pairs while the number of intervening items was increased. A key prediction was that, although the opportunity for memory trace short-term consolidation may protect against trace decay or inter-item interference, nevertheless reducing the number of intervening items within the ‘maintenance interval’ will be especially beneficial in an older age sample.

4.2 Method

4.2.1 Participants

Twenty-three older adults (18 female, M_{age} =77.74 years, age range = 67-93 years, $M_{education}$= 13.48 years, education range = 10-19 years) were recruited from local community centres in the Leeds area. All participants were native English speakers with normal or corrected to normal vision and were screened for neurological disorders and other medical condition thought to affect performance. One participant was excluded from analysis after disclosing they had previously suffered from a stroke (N= 22). Participants received £5 payment for their participation. Ethical approval was granted from the University of Leeds Ethics and Research Committee (ref: 14-0323).

4.2.2 Materials

Stimuli and equipment were the same as in Experiment 2.

4.2.3 Design and procedure

Participants completed the continuous recognition task, previously used in Experiment 2, arranged into 60-trial blocks, consisting of 20 single 'new' items, 10 'old' item pairs (an item presented and later presented again), and 10 'similar' item pairs (an item presented and a very similar item presented later). In a sequence of items participants reported by button press whether it was new (first time presented), old (the item had been previously presented) or similar (item shared visual features, but was not identical to, an item previously presented). A within subjects design was employed, with participants completing 3 experimental blocks for each lag condition, 4_long (4 intervening items with a 9.7 s III and 56.5s IRI), 12_short (12 intervening items with a 2.5 s III and 56.5s IRI), and 4_short (4 intervening items with a 2.5 s III and 12.5s IRI).

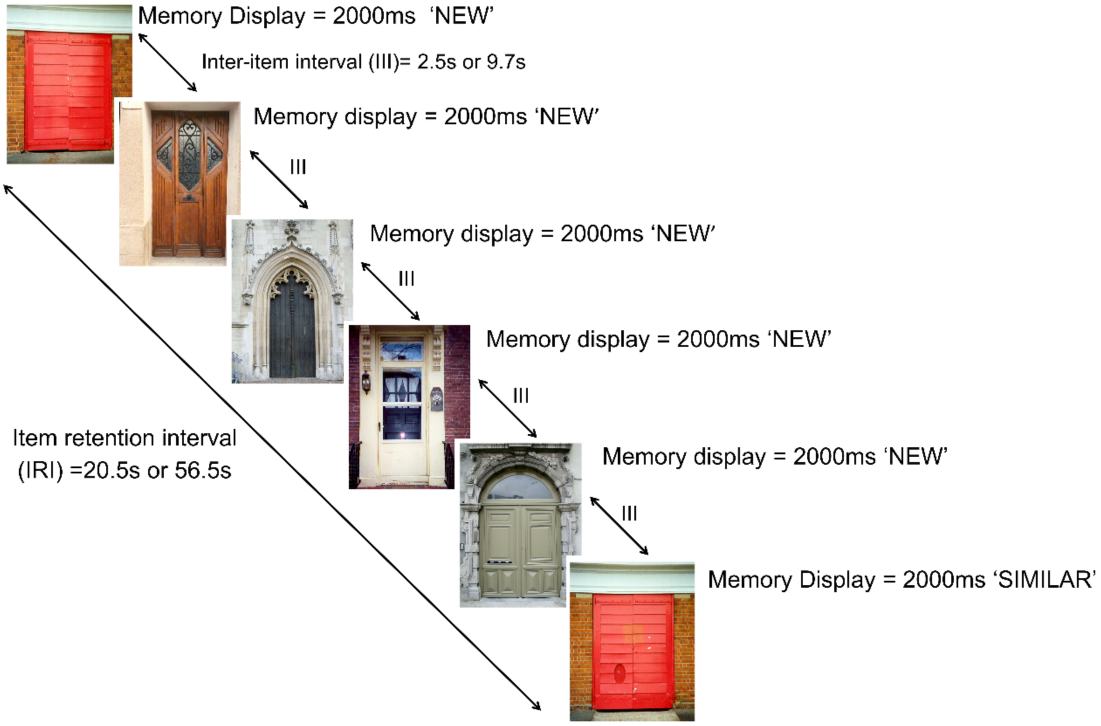

Figure 4-1. Example of the experimental procedure in the lag conditions: 4_long and 4_short.

As shown in Figure 4.1, all aspects of the trials mirrored Experiment 2 barring the following exceptions for the procedure. Participants completed 9 experimental blocks, divided across two test sessions. Unlike experiment 2, participants completed both sessions on the same day, with sessions separated by a minimum break of one hour. Participants completed two sessions with the order counterbalanced, as was the order of condition within a session. There was a 10 minute break between blocks and a minimum of one hours break between sessions. Visual items were not repeated across experimental blocks and participants were informed that items were not repeated from earlier blocks or sessions and as such their judgments of old, new and similar should be based on the current experimental block. Also, there was no significant difference in perceptual similarity between pairs used across our experimental conditions (χ^2 (12, *N* = 90) = 9.5, *p* = .66).

Prior to the continuous recognition task, participants completed a brief screening test thought to detect mild cognitive impairment. The Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005) is a brief validated 30 item screening tool which assesses a number of domains of cognitive function: visuospatial skills, executive function, memory, attention, language, conceptual thinking and orientation. The MoCA took approximately 10 minutes to complete.

Each participant was assigned a MoCA score out of a possible total of 30 and ideally, participants should have scored above 25. The Mean MoCA score was 25.91 (*SD* = 3.34). Fourteen participants scored 26 and above (high MoCA score; (M_{MoCA}=28.07, SD_{MoCA}=1.33, M_{age}=73.14, SD_{age}=6.19, $M_{\text{years of education}}$=14.57, $SD_{years\ of\ education}$ = 3.08) and eight individuals scored below 24 (low MoCA Score; M_{MoCA} = 22.13, SD_{MoCA} = 2.10, M_{age} = 84.63 years, SD_{age} = 6.70 years, $M_{years\ of\ education}$ = 12.00, $SD_{years\ of\ education}$ = 1.07). Preliminary analysis investigated any potential differences between the MoCA scores and performance on the continuous recognition task, in terms of the mean proportion of responses and the main outcome measure: RAS and BPS. A 2 (MoCA Score: high and low) x 3 (lag: 4_long, 12_short, 4_short) mixed ANOVA revealed no main effect of MoCA score for the RAS, *F (*1, 20) = .36, *p* = .554, η_p^2= .02 ; and BPS, *F (*1, 19*)* = .19, *p* = .665, η_p^2= .01. There was no interaction between the MoCA Score and lag conditions for the RAS, *F (*2, 40) = 1.20, *p* = .313, η_p^2= .056 and BPS, *F (*2, 38*)* = .16, *p* = .857, η_p^2= .008. Based on this, the eight individuals with a MoCA score of lower than 24 were included in the main analysis (See Appendix A for the full analysis of MoCA and outcome measures).

Table 4-1.

Definitions of the response possibilities for the new, old and similar stimuli; and formulae used for the main derived measures.

Stimulus	Response		
	New	Old	Similar
New	Correct rejection Rate	False alarm rate	Similar bias rate
Old	Miss rate	Hit rate	Incorrect
Similar	Incorrect	Pattern completion rate	Pattern separation rate

Recognition Accuracy Score (RAS) = Z(hit rate) – Z(false alarm rate)
Behavioural pattern separation (BPS) = Z(pattern separation rate) – Z(similar bias rate)
Behavioural pattern completion (BPC) = Z(pattern completion rate) – Z(false alarm rate)

4.3 Results

4.3.1 Overview of measures

The three outcome measures were calculated as shown again as a reminder in Table 4.1: the recognition accuracy score (RAS), the behavioural pattern separation score (BPS) and the behavioural pattern completion (BPC). As shown in Table 4.2, the mean raw proportion of responses was calculated for each stimulus condition per lag (4_long, 12_short, 4_short)[4]. Analysis was performed to check whether lag affected participants' ability to correctly identify novel items (correct 'new' responses); the mean proportions of correct rejections was calculated at each lag. A one way repeated

---

[4] Main outcome measures were calculated using a combination of response when the item was present and during the III. However, responses given during the III were only recorded if no response was provided when the item was present on the screen. This is to control for participants who had a slower response time than 2000ms or were inexperienced when using a computer keyboard.

measures ANOVA with factor of lag was not significant, *F (*2, 42) =1.75, *p* = .186, η_p^2= .08.

4.3.2 The RAS

For RAS, a one way repeated measures ANOVA with factor of lag was significant, *F (*2, 42) =7.25, *p* = .002, η_p^2= .26. Paired sample *t* tests comparing the separate effects of inter-item interval and number of intervening items, confirmed marginally higher performance at 4_long (*M* = 1.55, *SD* = .97) than for 12_short (*M* =1.14, *SD* = .94), *t* (21) = 3.51, *p* = .002, *d*= .44; and 4_short (*M* = 1.22, *SD* = .82); *t* (21) = 2.95, *p* = .008, *d* = .37. RAS for 4_short was not higher than for 12_short, *t* (21) = .73, *p* = .471.

Table 4-2.

Mean proportions (p) of response probability (new, old and similar) for each item type (New, Old and Similar) across lag (4_long, 12_short and 4_short). Standard deviation is given in parentheses (n=22).

Stimulus		Lag		
		4_long	12_short	4_short
New	Correct rejection rate	.59 (.23)	.54 (.24)	.55 (.25)
	False alarm rate	.18 (.12)	.20 (.16)	.19 (.16)
	Similar bias rate	.23 (.16)	.26 (.16)	.26 (.20)
Old	Miss rate	.12 (.13)	.16 (.15)	.16 (.15)
	Hit rate	.66 (.20)	.53 (.22)	.57 (.19)
	Incorrect	.22 (.15)	.31 (.21)	.27 (.17)
Similar	Incorrect	.20 (.16)	.28 (.14)	.32 (.18)
	Pattern completion rate	.46 (.15)	.41 (.17)	.40 (.15)
	Pattern separation rate	.34 (.15)	.30 (.14)	.28 (.18)

4.3.3 The BPS

For BPS, a one way repeated measures ANOVA with factor of lag (4_long, 12_short, 4_short) was significant, *F (*2, 40) = 4.86, *p* = .013, η_p^2 = .20. Paired sample t test confirmed performance was marginally higher for 4_long than for 12-short; *t* (21) = 2.68, *p* = .014, *d* = .63; and 4_short (*M* = .13, *SD* = .48); *t* (21) = 3.49, *p* = .002, *d* = .67. BPS for 12_short did not differ to 4_short, *t* (20) = .16, *p* = .872.

4.3.4 The BPC

For BPC a one-way repeated measures ANOVA with factor of lag (4_long, 12_short, 4_short) was not significant, $F\ (2, 42) = 1.80$, $p = .178$.

4.3.5 Comparison of performance for younger adults and older adults

To examine the effect of age on performance, each derived measure (RAS, BPS and BPC) as a function of lag was compared between older adult participants in the present experiment and the younger adults from experiment 2 ($N = 34$, $M_{age} = 19.56$ years). Separate 2 (age group: younger adult, older adult) x 3 (lag: 4_long, 12_short, 4_short) mixed ANOVA revealed no interaction between lag and age group for the RAS, $F\ (2, 108) = 1.81$, $p = .169$; BPS, $F\ (2, 102) = .94$, $p = .393$ and BPC, $F\ (2, 102) = .91$, $p = .41$. As shown in Table 3.3, there was a main effect of age group for the RAS, $F\ (1,54) = 25.10$, $p < .001$, $\eta_p^2 = .32$, BPS, $F\ (1, 51) = 10.99$, $p = .002$, $\eta_p^2 = .18$; and BPC, $F\ (1, 54) = 20.11$, $p < .001$, $\eta_p^2 = .27$.

Table 4-3.

Mean RAS, BPS and BPC for younger adults (n=34) and older adults (n=22). Standard deviation in parentheses.

		Lag		
		4_long	12_short	4_short
Recognition accuracy score (RAS)	Younger Adult	2.64 (.78)	1.93 (.74)	2.28 (.64)
	Older Adult	1.55 (.97)	1.14 (.94)	1.22 (.82)
Behavioural pattern separation (BPS)	Younger Adult	.76 (.55)	.46 (.41)	.60 (.49)
	Older Adult	.46 (.45)	.19 (.47)	.13 (.48)
Behavioural pattern completion (BPC)	Younger Adult	1.48 (.34)	1.21 (.50)	1.15 (.41)
	Older Adult	.91 (.58)	.79 (.59)	.72 (.44)

Separate independent t tests were conducted to compare each derived measure for younger and older adults within each lag condition. The RAS for younger adults was higher than older adults for the lag conditions, 4_long, *t* (54) = 4.65, *p* < .001, *d* = 1.25; 12_short, *t* (54) = 3.53, *p* < .001, *d* = .94; and 4_short, *t* (54) = 5.40, *p* < .001, *d* = 1.45. A similar age difference was identified in the BPS. The BPS for younger adults was marginally higher than older adults for the lag conditions, 4_long, *t* (53) = 2.15, *p* = .036, *d* = .60; 12_short, *t* (54) = 2.27, *p* = .027, *d* = .61; and significantly higher for

4_short, t (52) = 3.49, p = .001, d = .97. The BPC was higher for younger adults than older adults for the lag condition, 4_long, t (54) = 4.63, p < .001, d = 1.24; marginally higher for 12_short, t (54) = 2.82, p = .007, d = .77; and 4_short, t (54) = 3.67, p = .001 , d = 1.01. In summary, performance for younger adults was higher than older adults for the three measures. Finally, interest may also be given to the nature of the errors in the older sample, shown in Table 4.4: errors to 'new' items are noticeably higher than the younger sample, both in reporting new as similar (similar bias rate) and new as old (false alarms).

Table 4-4.

Contrast of older adult (n = 22) and younger adults (n = 34) for the main proportion of responses within each lag condition (4_long, 12_short and 4_short).

		Older adults		Younger adults				95% CI		
		M	*SD*	*M*	*SD*	*t* (54)	*P*	Lower	Upper	*d*
Correct rejection rate	4_long	0.59	0.23	0.82	0.12	-4.76	0.00	-0.32	-0.13	-1.31
	12_short	0.54	0.24	0.78	0.14	-4.64	0.00	-0.34	-0.14	-1.26
	4_short	0.55	0.25	0.79	0.16	-4.49	0.00	-0.35	-0.14	-1.17
Hit rate	4_long	0.66	0.2	0.79	0.17	-2.44	0.02	-0.22	-0.02	-0.7
	12_short	0.53	0.22	0.64	0.22	-1.75	0.09	-0.23	0.02	-0.5
	4_short	0.57	0.19	0.71	0.21	-2.56	0.01	-0.25	-0.03	-0.7
False alarm rate	4_long	0.18	0.12	0.07	0.06	4.8	0.00	0.07	0.16	1.22
	12_short	0.2	0.16	0.09	0.08	3.5	0.00	0.05	0.17	0.92
	4_short	0.19	0.16	0.07	0.05	4.29	0.00	0.07	0.18	1.14

		Older adults		Younger adults				95% CI		
		M	*SD*	*M*	*SD*	*t* (54)	*P*	Lower	Upper	*d*
Pattern Separation rate	4_long	0.34	0.15	0.27	0.14	1.71	0.09	-0.01	0.15	0.48
	12_short	0.3	0.14	0.22	0.13	2.25	0.03	0.01	0.15	0.59
	4_short	0.28	0.18	0.25	0.15	0.76	0.45	-0.06	0.12	0.18
Similar bias rate	4_long	0.23	0.16	0.12	0.1	3.1	0.00	0.04	0.18	0.85
	12_short	0.26	0.16	0.13	0.11	3.6	0.00	0.06	0.2	0.96
	4_short	0.26	0.2	0.14	0.13	2.74	0.01	0.03	0.21	0.73
Pattern completion rate	4_long	0.46	0.15	0.44	0.16	0.53	0.60	-0.06	0.11	0.13
	12_short	0.41	0.17	0.39	0.19	0.54	0.59	-0.07	0.13	0.11
	4_short	0.4	0.15	0.34	0.16	1.34	0.19	-0.03	0.14	0.39

4.4 Discussion

4.4.1 Summary of key findings

The aim of Experiment 3 was to examine performance of healthy older adults on the three-response recognition task designed to dissociate the influence of time interval and interference. Replicating the results of Experiments 1 and 2, performance benefitted when the post-item interval between successive items was extended. Note again that the post-item interval manipulation maintained an equal number of intervening items in two conditions' comparisons; when this interval was 2.5s (as is the case in conditions 12_short and 4_short), performance remained approximately the same across the measures despite both the number of intervening items and memory retention interval varying. This outcome strengthens confidence that post-item interval may be a fundamental controlling variable in this continuous recognition task, a phenomenon not previously reported. As already discussed, one explanation for benefits of extending the post-item interval in this task appeals to a particular retrieval strategy – recall to reject (Bakker et al., 2008; Kirwan & Stark, 2007; Lacy et al., 2011; Yassa et al., 2011). Consider the difficulty 'similar' items pose – first, the old item memory trace must be brought to mind, and then a comparison process and identification of the defining feature of the presented similar item must be realised. Surely this strategy must be made more demanding as we increase the similarity between old and similar items – or conversely reducing the salience of the defining feature difference (Yonelinas, 2002). Kim and Yassa (2013) report that the use of a recall to reject strategy, which was specified in their instructions to their participants, was associated with an increase in false alarms when stimuli were highly similar. And indeed this is what we observe in the older, relative to the younger, participants – somewhat higher false alarms.

4.4.2 Use of gist-based responses in older adults

It has been suggested that recognition in older adults may increasingly come to rely on gist based memory representations (Ally et al., 2013) – that is, signals of familiarity. Indeed a recent meta-analysis of recollection versus familiarity processes in recognition memory (Koen & Yonelinas, 2014) found that healthy aging is associated with a significant reduction in recollection with a medium to large effect size. Instead, recognition memory decisions in normal aging come to depend more on general signals of familiarity (Gallo, Sullivan, Daffner, Schacter, & Budson, 2004). Similarly, rather than recollecting the exact memory trace of a studied item, Paige, Cassidy, Schacter & Gutchess (2016) suggest that older adults rely on gist memory, which can result in false recognition when new and old items share similar perceptual features. If the older age participants do rely on gist based responding, it might not be surprising to observe an age difference in the task in the outcome measure of ‘old’ responses to new items (high false alarms) but not in reports of ‘old’ to old items– since old responses to old items will be accurately identified based on familiarity. So the older age sample may appear to be ‘as good as’ the younger sample in identifying old items, but rely on different information (gist-based versus precise recollection respectively).

However, it is worth noting that different retrieval strategies is only one possible explanation for the difference between behavioural pattern separation and pattern completion between the two age groups. Consider the role of item consolidation. What we might suppose, is that allowing time for item consolidation would be expected to offset the difficulty an older age sample will experience in making the more demanding similar response. That is, as the presumed fidelity of the memory trace for an item is made more robust through consolidation, the reliance on a gist-based response may make way for one based on item recollection – and possible ‘kick off’ a recall-to-reject strategy. The data reported in this chapter encourage this idea – similar responding does improve in older participants when post-item interval increases. In the next chapter, the aim will be to follow up these findings to establish whether the increased inter-item interval does indeed facilitate performance by providing time for item consolidation - or is there a possible role to be played by increases in the temporal distinctiveness of items?

Chapter 5 The question of time and memory consolidation

This chapter reports a further, novel investigation of item spacing within the three-response task. Further evidence is presented that extending time between items – and in particular the post-item interval - improves visual recognition memory. In contrast to the previous experiments, within the continuous sequence of door stimuli, item pairs were separated by four intervening items, generating a fixed delay between the first presentation of an item and its subsequent repeated presentation or its similar counterpart. Given that the focus was on available time for memory consolidation, a new manipulation was used here – the introduction of a post-item secondary task in the memory retention interval. Now, the work of Jolicœur & Dell'Acqua (1998) and of Nieuwenstein & Wyble (2014), has been influential in revealing the interplay between primary item encoding and secondary task load in this type of situation. As explained in Introduction, Section 1.7, Jolicœur & Dell'Acqua demonstrated that response times become slower in a secondary task inside a retention interval when only a brief post-item interval is allowed. They reasoned that during the brief interval, resources remain focused on consolidation of the memorandum, reducing the response time for the secondary task. More recently, the 'flip side' of this argument was demonstrated by Nieuwenstein & Wyble – these authors showed that a particularly demanding secondary task impaired primary task item encoding. Therefore these studies provide converging evidence that extended time for consolidation of a memory item can protect it from on-going interference. Before outlining the present experiment, it is worth noting that the inclusion of a consolidation mechanism into models of forgetting remains controversial and some say, confusing. For instance, Lewandowsky, Ecker, Farrell, & Brown (2012) refer to consolidation as an 'invisible' construct assumed to operate in the absence of forgetting but rarely demonstrated to occur. It is true that often, studies will attribute the lack of forgetting over an extended time interval to consolidation, without any direct manipulation of the process (e.g. Ricker & Cowan, 2014). When short term consolidation is directly manipulated, it is unclear, despite a few attempts such as Bayliss et al. (2015), as to whether short-term consolidation for non-verbal memory works across one or two seconds or several seconds, and whether there is justification for making a distinction between forms of consolidation and forms of maintenance such as ‘attentional refreshing’ (a somewhat puzzling concept appearing increasingly in the forgetting literature, having to do with maintaining a memory trace by thinking of it). These difficulties of interpretation will be revisited in Discussion. In the present experiment, a secondary pitch judgement task was introduced during the interval between each memory item. As discussed below, extending the window for consolidation increased accuracy on the secondary task, but more importantly, increased the accurate identification of old items. Yet the same outcome was not observed for similar items.

5.1 Experiment 4

The type of two task situation adopted by Jolicœur & Dell’Acqua (1998) is attractive for allowing manipulation of the time available for consolidation, whilst maintaining a fixed retention interval and level of interference between encoding and recall. To examine the benefit of an extended post encoding delay the present experiment introduced a similar two-task arrangement. Participants were required to complete the discrimination task making new/old/similar judgements but in addition a secondary task was introduced post-item. Again the stimuli were the picture of doors. In the interval between the door items a sequence of tones were sounded over headphones and participants indicated whether the tones were relatively ‘high’ or ‘low’ in pitch. The idea (supported by the research of Jolicœur & Dell'Acqua, 1998) was that this attentionally demanding pitch judgement would terminate ongoing item trace short-term consolidation – since attention is a 'bottleneck' (Bayliss et al., 2015; DeSchrijver & Barrouillet, 2017). That is, it is reasoned that participants cannot simultaneously direct resources to consolidation and to pitch judgement. Therefore, by manipulating the delay between door item offset and tone task onset the consolidation window would be fixed. As shown in Figure 5.1, the time available for consolidation was manipulated by varying the onset of the pitch judgement task, with two consolidation windows. The pitch judgement task was presented following either a short consolidation window of 2.5s or a long consolidation window of 9.7s. Note that in Nieuwenstein & Wyble (2014), the time for consolidation was limited to a duration of 250ms – 1,494ms. In the present situation, a longer interval was adopted, on the reasonable assumption that the three-response choice was more cognitively demanding than their one-response primary task. The prediction in the present experiment is straightforward – if benefits of extending the inter-item interval shown in Experiments 1 through 4 are indeed due to encoding benefits of short-term consolidation (and not, for example, simply reduced inter-item interference when the items are closer together in sequence), then the short consolidation window should impair performance (that is, impair encoding of items, presumably having consequences for all response types in the derived measures).

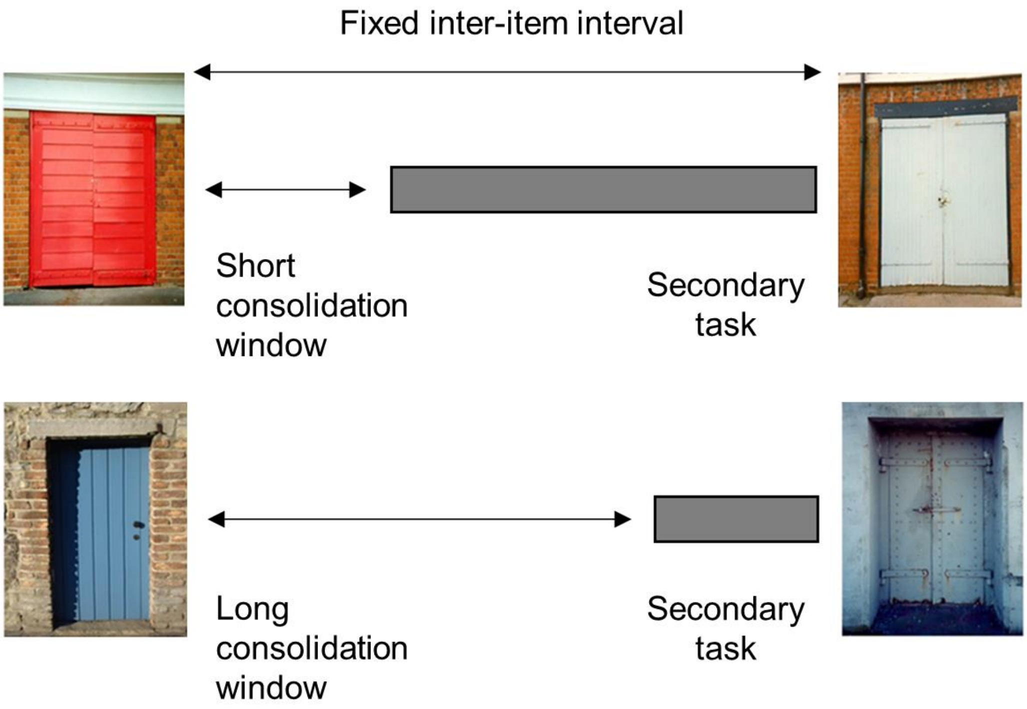

Figure 5-1. Illustration of the consolidation window conditions (short, long) controlled via the onset of a secondary task during the fixed inter-item interval between each item in the door task

5.2 Method

5.2.1 Participants

Thirty-five younger adults (29 females, M_{age}=19.34 years, SD_{age}=1.83 years, age range=18-29 years) were recruited from the School of Psychology at University of Leeds. All participants were native English speakers with normal or corrected to normal vision. Participants received course credit for their participation. Using previous effect sizes from study 2 and G*power analysis (Faul et al., 2007), a sample size of 33 was estimated to find a medium effect size (f= .21) with α= .05, 1-β= .80 and a moderate correlation between repeated measures of r= .50. Ethical approval was granted by the University of Leeds ethics and research committee (Ref: 15-0245)

5.2.2 Materials

The task used the same visual stimuli and apparatus arrangement as Experiment 2, except for the following changes. From the visual items employed in Experiment 2 and 3, 68 similar item pairs and 204 individual items were used as ‘new’ and ‘old’ items. Following each item, participants completed a pitch judgement Each tone presented lasted 100ms and was either high or low in pitch. Two pure tones with a frequency of 400 or 1200Hz, were generated using the software Audacity 2.0.3 (http://audacityteam.org) and presented through STAX SR-303 Classic headphones at a clearly audible level.

5.2.3 Design and procedure

Participants experienced a two-task arrangement, with the door task and a pitch judgement task alternating. The door identification task was the same as the previous experiment except for the following changes to the block length and the stimulus duration. The task was arranged into blocks of 30 trials, consisting of 10 single ‘new’ items, 5 ‘old’ item pairs (an item presented and later presented again), and 5 ‘similar’ item pairs (an item presented and a very similar item presented later). As shown in Figure 5.2., each trial began with a fixation cross centred on screen for 1000ms followed by the presentation of a visual item for 4000ms. As soon as the item appeared on screen, participants were instructed to indicate whether it was new (presented for the first time), old (presented previously in the sequence) or similar (similar but not identical to a previously presented item). Response were recorded using the keyboard keys ‘1’ for new, ‘2’ for old and ‘3’ for similar. The offset to onset of each item was separated by a fixed delay of 10s. During the interval at various stimulus-onset asynchronies (SOAs), participants completed a pitch judgement task. Participants were presented with a number of tones that were either high or low in pitch. Each tone lasted

100ms, following which participants indicated whether the tone was high or low pitch. Responses were recorded using the high and low labelled keyboard buttons. Once a response was provided, another tone was presented. The pitch judgment task was terminated at the end of the fixed delay, followed by the onset of the next door item.

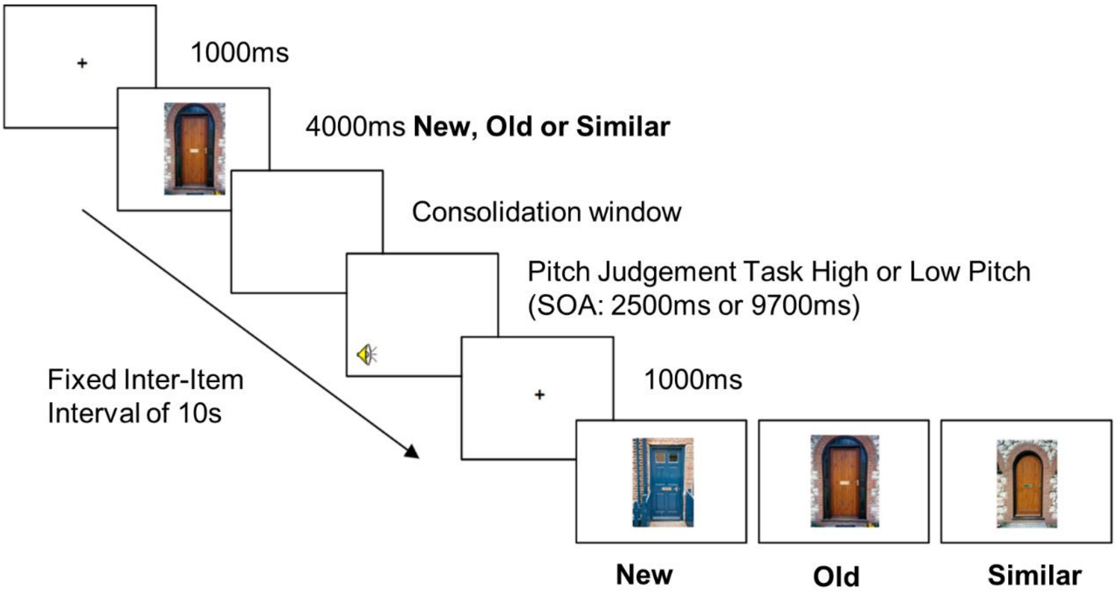

*Figure 5-2.*Schematic of the two-task arrangement with an example of new, old and similar stimuli used in the door task. During a fixed inter-item interval of 10s, participants completed the pitch judgement task at two SOA (2500ms and 9700ms), creating two consolidation windows (short and long). Old and similar item pairs were separated by four intervening trials.

In order to control the delay between item pairs, old and similar pairs were separated by 4 intervening items so that item pairs were separated by a fixed interval of 71s. Each successive item was separated by a fixed inter-item interval of 10s. The crucial measure was the SOA of the pitch judgement task which was either 2500ms or 9700ms. By varying the SOA, two consolidation windows were generated, a short window of 2500ms and a long window of 9700ms. In this within subject design, each participant completed 4 experimental blocks per consolidation window condition (8 experimental blocks in counterbalanced order, in total).

Participants completed two sessions with the order counterbalanced, as was the order of condition within each session. There was a 5 minute break between blocks and a minimum of 24 hours between each session (*M*=4.44 days, *SD*=3.10 days). Door items were not repeated across experimental blocks or session. Participants were informed that items were not repeated from an earlier block or session and as such their judgments of old, new and similar should be based on the current experimental block. To control for any effect of perceptual similarity, the earlier similarity data (presented in Chapter 6) informed the selection of items to ensure the perceptual similarity within each block and session was controlled.

Table 5-1.

Definitions of the response possibilities for the new, old and similar stimuli; and formulae used for the main derived measures.

Stimulus	Response		
	New	Old	Similar
New	Correct rejection Rate	False alarm rate	Similar bias rate
Old	Miss rate	Hit rate	Incorrect
Similar	Incorrect	Pattern completion rate	Pattern separation rate

Recognition Accuracy Score (RAS) = Z(hit rate) – Z(false alarm rate)
Behavioural pattern separation (BPS) = Z(pattern separation rate) – Z(similar bias rate)
Behavioural pattern completion (BPC) = Z(pattern completion rate) – Z(false alarm rate)

5.3 Results

5.3.1 Overview of measures

For the door task, the mean proportion of responses were calculated for each consolidation window (SOA of 2500ms or 9700ms). As before, RAS, BPS and BPC were calculated (a reminder is provided in Table 5.1). The ‘raw’ proportions are shown for each condition in Table 5.2. Outliers were identified using the proportion of novel items corrected as new. A one-way repeated measures ANOVA with a factor of SOA (short and long) was non-significant (p = .269).

Table 5-2.

Mean Proportion(p) for each response type (New, Old and Similar) across the SOAs (short 2500ms and long 9700ms). Standard deviation shown in parentheses.

		Short SOA	Long SOA
New	Correct rejection rate	.83(.11)	.81 (.11)
	False alarm rate	.06 (.04)	.04 (.03)
	Similar bias rate	.17 (.10)	.16 (.10)
Old	Miss rate	.15 (.11)	.10 (.11)
	Hit rate	.77 (.14)	.81 (.12)
	Incorrect	.09 (.09)	.09 (.07)
Similar	Incorrect	.24 (.14)	.26 (.19)
	Pattern completion rate	.40 (.17)	.35 (.19)
	Pattern separation rate	.37 (.14)	.38 (.18)

5.3.2 The RAS

The raw scores are shown in Table 5.2 and the outcome for the RAS is shown in Table 5.3. The RAS was significantly higher for the long SOA than short SOA, $t(34) = 3.92$, $p < .001$, $d = .84$.

5.3.3 The BPS

The raw scores are shown in Table 5.2 and the outcome for the BPS is shown in Table 5.3. SOA conditions did not differ by this measure, $t(34) = 1.58$, $p = .124$.

5.3.4 The BPC

The raw scores are shown in Table 5.2 and the outcome for the BPC is shown in Table 5.3. SOA conditions did not differ for this measure, t (34) = .51, p = .613.

Table 5-3.

RAS, BPS and BPC for Experiment 4 as a function of SOA (short, long). Standard deviation is given in parentheses.

	Short SOA	Long SOA
Recognition accuracy score	2.46 (.44)	2.83 (.44)
Behavioural pattern separation	.97 (.39)	.80 (.63)
Behavioural pattern completion	1.37 (.40)	1.42 (.48)

5.3.5 Pitch judgement task

During the fixed inter-item interval between successive items, a pitch detection task was presented following either a short or long consolidation window. Each tone presented was either high or low pitch. Performance for the tone task was examined using the accuracy and response time for the first tone presented (See Appendix B for analysis of the response accuracy for all the tones presented). As shown in Table 5.4, mean response accuracy (%) and reaction time (ms) for correct responses were calculated to assess task performance as a function of SOA. For mean response accuracy, a 2 (tone; high, low) x 2 (SOA: short, long) repeated measures ANOVA revealed an effect of SOA, F (1, 34) = 4.85, p = .035, η_p^2 = .13. Accurate identification of the first tone presented was marginally higher following a long SOA than short SOA, t(34) = 2.14, p = .040, d = .24. For reaction time, the same analysis did not reveal an effect of SOA, F (1, 34) = .24, p = .626; tone, F (1, 34) = .68, p = .415 or interaction, *F (*1, 34) = .23, *p* = .632. In summary, for accuracy but not reaction time, there was a modest suggestion of better performance in the pitch task when the SOA was long.

Table 5-4.

Mean response accuracy and reaction time in the pitch judgement task in Experiment 4 as a function of SOA (long, short) and tone pitch (low, high). Standard deviation is given in parentheses.

		SOA	
	Tone pitch	Long	Short
Response accuracy (%)	Low	92.03 (9.00)	90.36 (9.10)
	High	91.62 (9.03)	89.16 (8.57)
	Overall	91.76 (8.52)	89.74 (8.06)
Reaction time (ms)	Low	617.86 (95.19)	623.08 (94.32)
	High	611.83 (94.45)	613.64 (86.35)
	Overall	564.98 (62.15)	565.76 (61.98)

5.4 Discussion

5.4.1 Summary of results

Experiment 4 tests the idea that extended post-item time enhances pattern separation and completion performance. Two key temporal intervals were fixed, namely the inter-item interval between each item, and the retention interval separating item pairs. In this experimental arrangement, the goal was to terminate short-term memory trace short-term consolidation by introducing an attentionally-demanding pitch judgement task post-item. The interest was in the response profile for old, new, and similar recognition scores and whether these benefited from a relatively extended consolidation interval (keeping in mind the number of intervening items between old and similar item pairs was a constant). The outcomes were broadly consistent with the consolidation prediction - participants' ability to accurately recognise an old item was enhanced by increasing the window for consolidation. The RAS measure showed improved performance as the memory trace consolidation interval was increased from 2.5s to 9.7s. However, in contrast to the preceding experiments, the time available for consolidation had no significant effect in the other derived measures, BPS and BPC.

5.4.2 The effect of consolidation on accurate recognition memory

As explained in Introduction Section 1.6, recent investigations in visual memory have established that memory trace consolidation is a time-consuming process, and is reliant on a central attentional mechanism (Bayliss et al., 2015; Jolicœur & Dell'Acqua, 1998). Both the present findings and others employing a dual task paradigm, (Bayliss et al., 2015; Jolicœur & Dell'Acqua, 1998; Nieuwenstein & Wyble, 2014; Stevanovski and Jolicœur, 2007, 2011) have captured this idea by demonstrating that performance on both primary and secondary tasks is impaired by close temporal proximity (Nieuwenstein & Wyble, 2014). A similar pattern of results was obtained in Experiment 4, though the outcome in the secondary task was modest at best (a marginal improvement in accuracy when the post-item interval was extended). One notable feature of the work of Jolicœur and Dell'Aqua (1998), is that they failed to demonstrate a consolidation benefit in their primary test of memory. Yes, the longer interval permitted their participants to perform a little faster on their secondary task, but memory performance in the primary task was unaffected. Of course, they used very different stimuli and intervals to the ones used here. Some (e.g. Ricker, 2015) claim a process of 'short-term' consolidation is complete within a couple of seconds, but they are quite vague as to the experimental justifications for such a claim. Others, such as Bayliss et al. (2015), have tried to manipulate consolidation windows of the order of a couple of seconds too. The present decision to study longer temporal windows was, as outlined in the introduction to this chapter, driven by two factors. Firstly, Experiment 2 and 3 identified benefits of extending the free interval between items in the primary task well beyond two seconds. Secondly, the non-verbal stimuli also encourage one to explore temporal parameters well beyond those popular in the study of verbal items. For example, the work of McKeown and Mercer (2012) has been influential in encouraging workers in non-verbal memory to look beyond the few-second window of contemporary verbal memory research.

However, given the duration of the present consolidation window, it is perfectly reasonable to ‘sit on the fence’ as to what the post-item window permits: be it consolidating brief traces into the short-term memory, or permitting the refreshing or maintenance of those items, protecting them against inevitable ‘decay’. Both consolidation and maintenance processes such as attentional refreshing and articulatory rehearsal are believed to operate following the encoding of a stimulus (Bayliss et al., 2015). In regards to rehearsal, an argument has already been put forward in Introduction as to how employing a large number of within category visual stimuli eliminates or severely limits the use of verbal labelling and rehearsal. A more interesting question that needs addressing is, given their reliance on attentional resources, are short term consolidation and attentional refreshing separate processes? Within the context of the time-based resource sharing model of forgetting, attentional refreshing is typically manipulated by varying the cognitive load. Frequently, cognitive load is defined according to the proportion of time an intervening secondary task occupies attention away from the memorandum. Based on their manipulation of cognitive load and consolidation, both Bayliss et al. (2015) and De Schrijver &

Barrouillet (2017) concluded that consolidation is rapidly completed before maintenance processes begin. Yet, despite the growing body of evidence in support of short term consolidation, it is absent in contemporary models of forgetting.

5.4.3 Conclusion

In summary, once again the duration of the post-item interval within the three-response task has been identified as a key factor determining the accurate retention and recognition of a previously stored memory items. Accurate recognition of a repeated item was moderately improved by extending the temporal interval between the visual item and the onset of a secondary task. Both performance on the visual task and the secondary auditory task declined slightly when presented in close temporal proximity. However, the effect was not identified in the response profile for similar items. Thus, the remaining chapters shift the focus towards the effect of *interference* on accurate identification of similar items. So the role of mnemonic similarity between similar item pairs is examined in the next experiment.

Chapter 6 The question of stimulus similarity

This chapter further develops the over-riding question in this thesis – how to develop a more valid test of the putative memory mechanisms of study? As explained in Introduction, Section 1.7, a critical governing variable governing behavioural recognition performance for immediate memory is the physical closeness within feature-space of presented stimulus items. Actually, one of the reported advantages of the BPS task is that it appears to be highly sensitive to changes in ‘perceptual’ similarity between item pairs (Stark et al., 2013). Attempts have been made to vary stimulus similarity parametrically for BPS stimulus sets. For example, in Yassa, Lacy et al. (2010) each *similar* item pair was assigned a measure of perceptual similarity. This similarity metric was based on both subjective ratings of perceptual similarity and recognition performance, specifically the proportion of *incorrect old responses*. Based on the similarity metric, item pairs were allocated into discrete groups or ‘bins’ ranging from most similar (bin 1) to least similar (bin 5). Interestingly, their bias metrics included two of the derived measures reported throughout this thesis: BPS and BPC, but also an additional one: RAS. Crucially, tests of recognition memory performance by these derived measures mapped onto the similarity bins. Their stimuli were the familiar cartoon-like pictures of everyday objects. In this chapter similarity ratings are used to develop a set of the door stimuli varying parametrically. In Experiment 5a, a set of door stimuli varying in perceptual similarity is developed using ratings by participants of 90 similar item pairs. Then, in Experiment 5b, a fresh group of participants evaluates the new stimuli in a yes/no variant of the BPS. Interest was two-fold: how would the new stimulus set map onto performance in the old/new task (the responses to the stimuli were simply ‘old’ or ‘new’) where items varied across old, new and similar forms? (otherwise, do the new similar items map appropriately onto accurate new response, where a response ‘new’ to a similar item is taken as correct); secondly, does the new/old variant promise to be a more tractable test (for example, in interpretation of responding appropriately to variations in stimulus similarity, and in calculating response bias)? Finally, in Experiment 5c, selected data for younger and older participants in two of the previously reported experiments (Experiment 2 and Experiment 3 respectively) are re-analysed in light of the similarity rating distribution for similar items; by ‘selected’ the is meant the analysis only focuses on the similar response type.

6.1 Experiment 5a

6.1.1 Method

6.1.1.1 Participants

Twenty-nine participants (27 female), aged between 18 – 28 years (M_{age} = 18.97 years, SD_{age} = 2.04 years) were recruited from the School of Psychology at the University of Leeds. Participants were all native English speakers, with self-reported normal to corrected-to-normal vision. All participants were native English speakers with self-reported normal to corrected to normal vision. Participants received course credit upon completing this study. Ethical approval was granted by the University of Leeds ethics and research committee (reference -14-0179).

6.1.1.2 Materials

Visual stimuli consisted of 180 coloured photographs of doors arranged into 90 similar item pairs. These pairs, the same as used in Experiment 2 and 3, were matched in terms of categorical similarity according to: function, age, colour, glazing, condition, shape, opening, details, surround and memorability. Each image was adjusted to a height of 227 pixels and width of 178 pixels. Items pairs were presented individually, with the two items presented side by side. Visual stimuli were presented on a white background, on a Dell 1708FP monitor, with participants approximately 60cm away from the monitor. The experiment was run using E-Prime 2.0 software (Schneider et al., 2002). Example pairs are shown in Figure 6.1

Figure 6-1. Examples of similar door item pairs used in the Experiment 5a and 5b. Stimuli in the top row were rated high similarity and bottom row were rated low similarity.

6.1.1.3 Design and procedure

Participants rated the perceptual similarity of each item pair using a Likert scale of 1 to 7, with 1 representing ‘weak similarity’ and 7 representing ‘strong similarity’. As shown in Figure 6.2, each item pair was presented on screen one at a time with a reminder of the scale present. Participant responses were collected using the keyboard numbers 1 through 7. A fixation cross lasting 500ms was presented on screen prior to each item pair. The similarity rating task was self-paced and participants had as long as needed to provide their similarity ratings. On average, participants viewed each image pair for 4784.75ms (*SD* = 1154.53ms). The order of the pairs was randomised across participants and after each 10th trial, participants were provided with an opportunity for a break.

To encourage participants to use full rating scale, prior to the similarity ratings, participants viewed all 90 item pairs, individually. Participants were instructed to take their time to become familiar with the changing degree of similarity across the item pairs. On average, each item pair was viewed for 4446.07ms (*SD* = 3505.18ms). Excluding breaks, the average duration of the testing was 16.30 mins (*SD* = 4.71 mins, range = 9.08 mins - 29.32 mins)

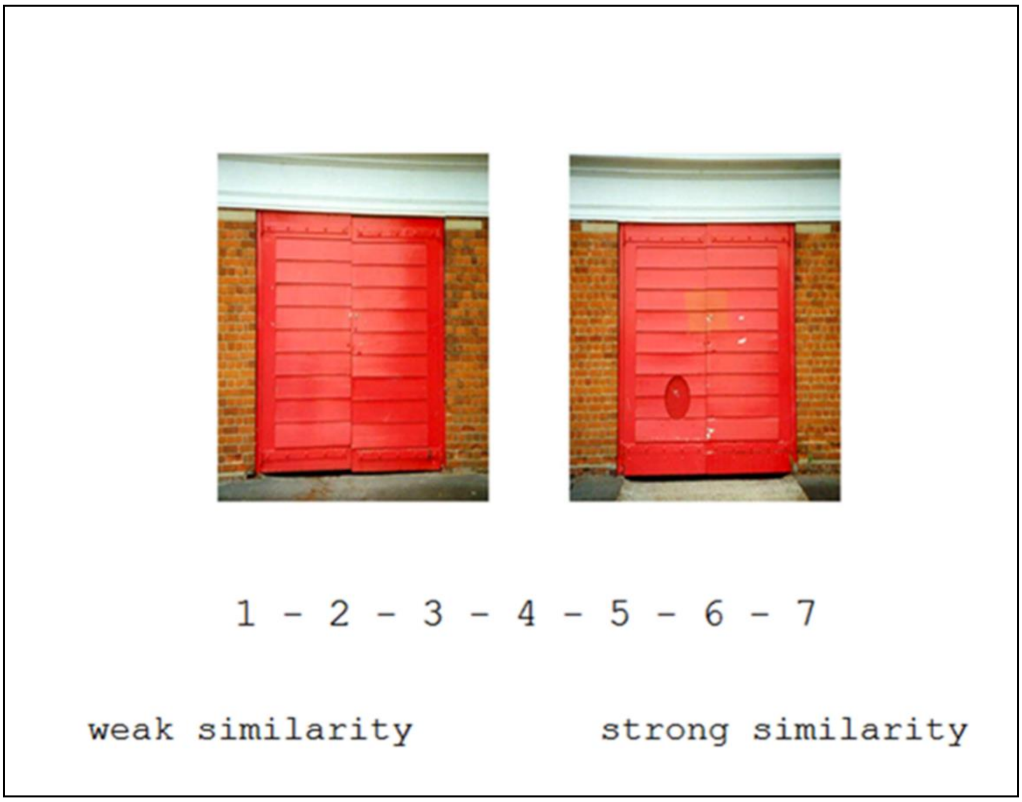

Figure 6-2. Example of an item pair and rating scale presented during the similarity rating task for a single trial.

6.1.2 Results

Ratings of perceptual similarity were averaged across participants. Each item pair was assigned a similarity rating of 1-7 based on the mean rating of perceptual similarity. Overall, across the 90 item pairs, the mean rating was 4.59, with a range of 1 to 7, and with a standard deviation of 1.70. Therefore we can be confident that the participants were using the full range of the scale. Table 6.1 outlines the number of item pairs for each similarity rating alongside the mean similarity rating and the Cronbach's alpha score for that scale. Additionally, each item pair was classified as either a low similarity pair or high similarity pair. Items with a rating of 1-4 were allocated to the low similarity group (*n* = 41). Item pairs with a rating of 5-7 were allocated to the high similarity group (*n* = 49). The mean similarity rating between the low similarity (*M* = 3.67, *SD* = .72) and high similarity group (*M* = 5.35, *SD* = .85) significantly differed, *t* (88) = 10.06, *p* < .001, *d* = 2.15.

Table 6-1.
Mean similarity rating for each similarity rating group; n is the number of item pairs per similarity rating. Standard deviation is given in parentheses.

Similarity Rating	Similarity Group	*n*	*M*	α
1	Low	3	2.28 (1.36)	.27
2		9	2.97 (1.37)	.68
3		16	3.91 (1.45)	.84
4		13	4.18 (1.43)	.74
5	High	20	4.64 (1.37)	.82
6		14	5.28 (1.29)	.81
7		15	6.37 (.98)	.65

6.2 Experiment 5b

In this second phase, the validity of the perceptual similarity ratings for the door items was examined in a new sample of younger adults using a *two response* behavioural discrimination task. In contrast to the previous studies, participants were required to provide a new/old response (see for example, Loiotile & Courtney, 2015 for a similar testing method). Previous research has established that the perceptual similarity between similar item pairs affects performance. Specifically, accuracy declines with increasing similarity between items within the sequence. The consequence is that proportion of correct responses to similar items decreases as perceptual similarity increases (Lacy et al., 2011; Pidgeon & Morcom, 2013; Yassa et al., 2011). Here it was predicted that if the similarity rating accurately reflects *perceptual similarity*, performance for similar items should become worse as the perceptual (previously established) similarity increases across the set of door stimuli from the first phase (Migo, Montaldi & Mayes, 2013). The focus therefore was simply on the *similar response profiles* across the judged similarity measure.

6.2.1 Method

6.2.1.1 Participants

Twenty-one participants (16 female), aged between 18-30 years old (M_{age} = 19.67 years, SD_{age} = 2.54 years) were recruited from the School of Psychology at the University of Leeds. Participants were all native English speakers, with self-reported normal to corrected-to-normal vision. Participants received course credit upon completion. Ethical approval was granted by the University of Leeds ethics and research committee (ref: 16-0050).

6.2.1.2 Stimuli and procedure

The task was arranged into experimental blocks of 60 trials, consisting of 20 single new items, 10 old item pairs and 10 similar item pairs. Each door item was presented one at a time for 4s, followed by an inter-item interval of 3s. Old and similar item pairs were separated by four intervening items with an item retention interval of 31s. Each experimental block lasted 7.05 minutes with a 10 minute break between each block. Participants were instructed that they would be presented with a continuous sequence of pictures of doors and each could be classified as ‘new’ (first time presented in the current sequence) or ‘old’ (item has been presented previously in the sequence). Using labelled keyboard response buttons, participants were instructed to decide for each item whether it was new or old. No similar response type was used here. For the two response continuous recognition task, the correct response to the presentation of a similar item is ‘new’ (Loiotile & Courtney, 2015). New, old and similar item pairs were randomly allocated into one of nine possible experimental blocks. Each participant completed three experimental blocks in one session. Selection of experimental blocks was counterbalanced across participant. Similar item pairs were randomly allocated in one of the nine experimental blocks. The allocation of similar item pairs from each similarity group was equally distributed across the blocks (χ^2 (48) = 59.39, p = .125) and degree of similarity group (χ^2 (8, N = 90) = 11.38, p = .181.

6.2.2 Results

6.2.2.1 Analysis of pattern separation rate and pattern completion rate

Table 6.2 shows the distribution of the 2-type (new/old) responses for the task across the three types of stimulus and Table 6.3 shows performance for similar responding using the pattern separation and pattern completion rate across similarity judgements. Recall that., for a two response continuous recognition task, pattern separation is the proportion *of similar items correctly identified* as ‘new’ and the pattern completion rate is the proportion of *similar items incorrectly identified* as ‘old’.

Table 6-2.

Mean proportions of responses (new and old) for each item type (New, Old and Similar). Standard deviation is given in parentheses (N = 21).

Stimulus		*M (SD)*
New	p(new)	.91 (.06)
	p(old)	.09 (.06)
Old	p(new)	.12 (.07)
	p(old)	.88 (.07)
Similar	p(new)	.49 (.15)
	p(old)	.51 (.15)

If similarity maps onto the two measures of interest here, one should go up with similarity and one should go down. A 2 x 2 mixed ANOVA with the repeated measures factor of response type (pattern separation rate, pattern completion rate) and a between subjects factor of similarity rating group (low similarity, high similarity) revealed a moderately significant interaction, F (1,88) = 4.60, p = .035, $\eta_p^2 = .05$. Independent t tests showed the pattern separation rate was marginally significantly higher for the low similarity group (M = .57, SD = .27) than the high similarity group (M = .45, SD = .25), t (88) = 2.15, p = .034, d = .23. For the pattern completion, the expected reverse pattern (but again, only marginally) was observed between low similarity group (M = .44, SD = .27) and high similarity group (M = .55, SD = .25), t (88) = -2.14, p = .035, d = .23. In summary, this first attempt to parametrically vary the new door stimuli to create an 'index of perceptual similarity' mapping onto discriminative performance was moderately successful.

Table 6-3.

Mean pattern separation rate (proportion of similar items correctly identified as old) and pattern completion rate (proportion of similar items incorrectly identified as new) as a function of perceptual similarity rating (where 7 is 'high'). Standard deviation is given in parentheses.

Similarity rating	Pattern separation rate	Pattern completion rate
1	.70 (.15)	.30 (.15)
2	.63 (.26)	.38 (.26)
3	.57 (.31)	.43 (.31)
4	.49 (.24)	.51 (.24)
5	.57 (.24)	.43 (.24)
6	.30 (.20)	.70 (.20)
7	.42 (.24)	.58 (.24)

6.2.2.2 Criterion bias for the new/old variant

In Introduction, Section 1.9 it was acknowledged that the three-response BPS presented considerable difficulties in interpretation of decision strategies by participants since responding spans two decision boundaries. The two-response new/old task may have binary choices but it too presents difficulties since the underlying stimulus distributions themselves span two decision boundaries – and with Macmillan and Creelman (2005) decision bias c is not calculable. As a (tentative) exploration for the two-response data here, however, an assumption was made that 'new' response to underlying similar stimulus types is 'correct'. So two criterion c's were calculated: -0.5 times z[hits (old responses to old items)] plus z[false alarms (old responses to new items)]; and -0.5 z[hits (new responses to similar items)] plus z[false alarms (new responses to old items)]. These calculations revealed the data shown in Table 6.4. A summary statement can be made for these tentative calculations. First, 8 of 21 participants appear slightly ‘liberal’ in responding old to old items; secondly, all participants appear relatively neutral or slightly conservative in reporting ‘similar items’ as new. Therefore the calculations match expectations if participants are indeed sensitive to the underlying changes to the similar items.

Table 6-4.

Two forms of criterion c measure (see text, Section 6.2.2.2) calculated for new/old data in Experiment 5c.

Participant	Old to Old C		New to Similar C	
	d’	c	d’	c
1	2.46	0.26	0.80	0.57
2	2.94	0.36	0.68	0.77
3	2.78	-0.11	1.67	0.67
4	1.87	-0.03	0.88	0.53
5	2.94	-0.36	1.49	1.09
6	2.93	0.18	1.20	0.68
7	2.66	0.05	1.71	0.43
8	3.14	0.07	0.88	1.06
9	2.45	0.60	0.62	0.31
10	2.88	-0.06	1.33	0.83
11	2.25	-0.16	0.55	1.00
12	1.59	0.09	0.92	0.24
13	3.34	0.17	1.16	0.92
14	2.60	0.33	0.97	0.48
15	2.74	-0.47	2.09	0.79
16	3.96	0.15	1.92	0.88
17	2.22	0.00	1.64	0.29
18	2.47	-0.27	1.25	0.88
19	2.49	0.40	0.84	0.42
20	2.80	0.43	1.93	0.00
21	3.34	-0.17	2.17	0.75

6.3 Experiment 5c

In this third phase, a reanalysis of the ageing data already collected is performed based on the new similarity ratings of the stimuli. The experimental tests reported above manipulated similarity and demonstrated that errors of ‘old’ were common when items were dimensionally ‘similar’ (as pre-determined in the ratings data). This was of course expected, and is perfectly in line with recent reports demonstrating the decline in pattern separation with increased mnemonic similarity between a target item and its similar counterpart (Lacy et al., 2011; Stark et al., 2013; Stark et al., 2015; Yassa et al., 2011). It has been established that older adults may be more susceptible to a failure to make a discrimination under conditions of increased mnemonic similarity (Stark et al.,2013; Yassa et al., 2011) and so in the analyses reported below, age differences in response to *similar item pairs* were examined as a function of perceptual similarity. Data have been reported already from two age groups. Therefore, in trying to better understand stimulus similarity on measures of behavioural pattern separation, a re-analysis of the data from Experiment 2 (younger adults) and Experiment 3 (older adults) was conducted. The re-analysis focused on how good the two age groups were in identifying similar items, so for simplicity (and clarity) just two measures are reported – one for accuracy (identifying similar items) which is termed ‘pattern separation rate’ – and one for errors (reporting similar items as old) which is termed ‘pattern completion rate’.

6.3.1 Method

The analyses reported were a re-examination of a sub-set of the data arising from Experiment 2 and Experiment 3 (only the similar response type was focused on). The total sample consisted of 34 younger adults (YA; 31 female; M_{age} = 19.56 years, SD_{age}

= .88) and 22 older adults (OA; 18 women, M_{age} = 77.74 years). In contrast to the two-response form of task reported in Experiment 5b, data in this re-analysis were based on experiments which collected three responses: new, old and similar.

6.3.2 Results

Based on the results of Experiment 5a, each similar item pair was assigned a similarity rating and allocated into a similarity group (low or high similarity ratings). This exploratory analysis just looked at the two responses to similar items – one for accuracy and one for errors (although it is acknowledge several other aspects of the data might legitimately be used). Table 6.5 respectively shows the two responses to similar items as a function of perceptual similarity across the two age groups (younger adults, older adults).

A 2 x 2 x 2 mixed ANOVA with repeated measure of response type (pattern separation rate, pattern completion rate), similarity (low, high), and an independent measure of age (younger adult, older adult) showed a main effect of response type, F (1,54) = 31.17, $p < .001$, $\eta_p^2 = .36$; main effect of similarity, F (1,54) = 16.16, $p < .001$, $\eta_p^2 = .23$; and interaction of similarity with response type, F (1,54) = 17.79, $p = .032$, $\eta_p^2 = .24$; and of similarity with age, F (1,54) = 6.32, $p = .015$, $\eta_p^2 = .103$. Pairwise comparisons between low and high similarity conditions revealed that participants' pattern separation rate declined as similarity increased (younger adults: t (33) =3.10, $p = .004$, d= .39; older adults; t(21) = 2.57, $p = .018$, $d = .39$), although only statistically marginally (given the adjustment to p for multiple comparisons)

For pattern completion rate, a 2 (similarity: low, high) x 2 (age: younger adults and older adults) mixed ANOVA revealed a main effect of similarity, F (1,54) = 11.51, $p = .001$, $\eta_p^2 = .17$, and marginally significant interaction, F (1,54) = 4.14, $p = .047$, $\eta_p^2 = .05$. Post hoc comparisons revealed that for older adults, errors did not differ with similarity, *t* (21) = .61, *p* = .548; whilst for younger adults, *errors differed significantly* for low and high similarity, *t*(33) = 7.97, *p* < .001, *d* = .94.

Table 6-5.
Mean pattern separation and pattern completion rate as a function of perceptual similarity (low similarity and high similarity) for younger adults and older adults. Standard Deviation is given in parentheses.

Response	Similarity	Younger adults	Older adults
Pattern separation rate	Low	.28 (.14)	.28 (.18)
	High	.23 (.11)	.22 (.15)
Pattern completion rate	Low	.30 (.14)	.43 (.21)
	High	.44 (.16)	.47 (.19)

6.4 Discussion

It is just worth reminding the reader that the three phases of Experiment 5 do different things with different forms of data. Phase 'a' has to do with ratings of the door stimuli, phase 'b' reports outcomes for a two-response task, and phase 'c' does not report new data, but is a re-analysis of selected already reported outcomes of the three-response task in light of the new assignment of the door stimuli according to mnemonic similarity.

As described in Introduction, three questions were posed: firstly, whether the new stimulus set maps onto performance in the yes/no task, where a response 'new' to a similar item is taken as correct; secondly, whether the yes/no variant promises to be a more tractable test (for example, in interpretation of responding appropriately to variations in stimulus similarity, and in calculating response bias); and finally, whether younger and older participants' responding (from Experiments 2 and 3) to similar items would differ along the new mnemonic similarity continuum established here. The outcomes were promising, if not fully realised. Notably, performance in the yes/no task did approximately map onto the 7-point similarity scale (but note the 'marginal' qualification in the t-tests in Experiment 5b). Regarding the re-analysis data, the focus was only on accuracy (pattern separation rate: accuracy of reporting similar to similar items) and errors (pattern completion rate: errors in reporting old to similar items). For accuracy, performance did map onto the similarity scale appropriately (but again, note the 'marginal' in the t-tests). For errors there was a statistically reliable and important contrast: whereas for the younger group errors were significantly higher as item similarity increased (thus they were showing sensitivity to feature overlap), for the older group the low and high similarity items were much of a 'sameness'.

This contrast is reminiscent of visual recognition failures identified in the literature. Most simply described, the hippocampus supports the visual memory system by binding details or features to form new associative memories during encoding, as well as retrieving and reconstructing these details upon recall (Paige et al., 2016; Yassa & Stark, 2011). While this constructive process may provide a reliable record of an ever changing visual environment, it is prone to errors (Gutchess & Schacter 2012). Firstly, we may fail to recognise previously presented information or secondly, newly encountered information could be mistakenly treated as though it has been previously encountered (Koutstaal & Schacter, 1997). Some refer to this as 'false recognition' and indeed, it is frequently observed in older adults (Koutstaal & Schacter, 1997; Pidgeon & Morcom, 2013; Yeung et al., 2013). Indeed this inability to discriminate between somewhat similar stimuli has been identified as the crucial feature defining memory impairments in older adults (Pidgeon & Morcom, 2013). Yeung et al (2013) monitored eye tracking behaviour as participants studied a sequence of everyday objects belonging to the same semantic category. Participants then viewed a sequence sharing either high or low perceptual similarity (sharing similar physical features) with items from the study phase. They found that in comparison to younger adults, mean eye fixations in older adults did not differ across old repeated items and items sharing high similarity; indicating that these highly similar items were falsely viewed as old. There was no age difference when viewing repeated or low similarity items. This contrasting pattern of behaviour has since been confirmed (Pidgeon & Morcom, 2013; Stark et al., 2013; Yassa, Mattfeld et al., 2011). Furthermore, by manipulating the perceptual similarity of item pairs, neuroimaging studies employing BPS have demonstrated changes in activity of DG/CA3 region in response to increased stimulus similarity. For dissimilar items, there was no significant difference between activity in the DG/CA3 between older and younger adults (Lacy et al., 2011; Yassa, Mattfeld et al., 2011); but in older adults, DG/CA3 activity declined with increasing similarity.

Yet to date, it has been unclear whether age related increases in such false recognition was driven by overlapping *semantic* information between memory items (Koutstaal, 2003), or as proposed by computational models of pattern separation, results from an impaired ability to mnemonically discriminate perceptually similar memory items (e.g. ; Stark et al., 2013; Wilson et al., 2006). It has been argued that increased false recognition with age may be attributed to an overreliance on gist based responding (Koutstaal, 2003; Koutstaal & Schacter, 1997). As we age, the encoding of visual information becomes less distinctive; this not only leads to the use of gist-based responding but a move away from the ‘percept’ leads to reliance on verbal labels (Brainerd & Reyna, 1990; Devitt & Schacter, 2016). In their semantic categorization account of false recognition, Koutstaal (2003) propose that increased false recognition in older adults reflects an emphasis on the semantic processing of visual items during encoding. According to this account, false recognition is driven by a shared semantic category, rather than perceptual similarity. Unfortunately, the majority of studies of pattern separation and ageing have employed everyday objects lending themselves to verbal or category labelling. The present findings with single-category stimuli support the view that perceptual, not semantic, features are responsible for the age contrast observed here.

Moving on, having established our ‘metric’ of similarity with the new stimulus set, it is possible to pursue theoretically interesting questions – and in the next and final chapter the question of inter-item interference in the memory task is examined. It is acknowledged, of course, that as we increase item similarity across successive presentations of the stimuli, it is creating precisely the situation where item-on-item feature interference may occur. Thus items may increasingly suffer from inter-item interference. The continuous recognition task is one demonstrating discrimination of differences between internal representations of items (the memory trace) and a current stimulus (the new, old or similar item), but it is also arguably a test that suffers from perturbation of the internal representation (the memory trace) produced by the current stimulus encoding. As we manipulate inter-item similarity it becomes obvious that we will be influencing both discrimination and perturbation. Indeed novel stimuli may command most attending and support most perturbation in the sense just described. That is, in studying participants’ ability to notice change or novelty in the stimuli, we may inadvertently be changing the salience of those stimuli trial by trial. The final experiment examines some of these possible inter-item effects, and asks how *item novelty* may be governing *inter-item interference*.

Chapter 7 The question of stimulus novelty

The aim of Experiment 6 was to pursue the question of inter-item interference. The starting point was to consider how the items themselves (new, old, and similar across successive trials) may differ: 1) in their vulnerability to proactive interference; and 2) in their effectiveness as retroactive interferers. One key reason for their differing effectiveness as interfering events is that they differ *in novelty*. The idea tested here is that highly novel items will be attentionally arresting and so interfere more. Therefore it was arranged that target item pairs (old and similar) were separated by four intervening items which varied in their degree of novelty. The outcomes were interesting. Firstly, the ability to accurately recognise an old target item declined most when the intervening items were four unique novel items. In contrast similar intervening items appeared to offer less interference. Therefore the findings reported here are the first to identify a critical factor in the continuous recognition task – successive item interference is governed by item novelty.

7.1 Experiment 6

The design was quite straightforward. Within the continuous sequence of items of doors, old and similar target item pairs were separated by 4 intervening items. By manipulating the degree of novelty across these four intervening items, it was intended to vary the degree of interference generated between the initial encoding of a target item and its subsequent repeat or similar item pair. The degree of interference was varied by manipulating the perceived novelty across the four intervening items, starting with four unique novel items and then reducing the degree of novelty, using four individual items which were similar to one another, and in a third condition the four successive items were the same. The interference generated by the intervening items was expected to influence both the accurate recognition of an old target item as well as the false recognition of similar target items. It was predicted that performance would decline as the novelty across the intervening items increased. Finally, a secondary aim was to examine age differences in performance as a function of novelty encoding of the intervening items. Based on age related changes to novelty detection in visual memory, it was predicted that degree of interference would interact with age. An interesting consideration is that, if the ability to identify novel features or novelty per se declines with age, an older aged sample might even show release from interference.

7.2 Method

7.2.1 Participants

Fifty one participants (42 females, M_{age} = 47.43 years, SD_{age} = 22.77 years, range = 18 – 89 years) were recruited from the University of Leeds community. Participants were all native English speaker with self-reported normal or corrected to normal vision and no prior diagnosis of any cognitive impairments. Based on their age, participants were divided into two age groups, younger adults and older adults. There were 28 younger adults, aged between 18-59 years (6 males, M_{age} = 29.00 years, SD_{age} = 11.45 years) and 23 older adults, aged between 60 – 89 years (3 males, *Mage* = 69.87 years, SD_{age} = 7.66 years). There was a significant difference in age between the groups (younger and Older), t (49) = 14.62, $p < .001$, d = 4.28. There was no significant difference in years of education between the younger (*M*=16.79, *SD*=2.99, range=12-23) and older adults (*M*=16.30 *SD*=3.14, range=7-21), t (49) = .56, p = .578, d = .16. The Montreal Cognitive Assessment Test (MoCA; Nasreddine et al., 2015) was administered to participants in the older adults to screen for possible cognitive impairment. All older adults participants obtained a score of 26 and above (M_{MoCA} = 28.

32, SD_{MoCA} = 1.36, range = 26-30 points). Ethical approval was granted by the University of Leeds ethics and research committee (ref: 16-0050). Participants received a monetary payment of £5 upon completing the study. Two younger adults were initially recruited but have been excluded from all data analysis due to an error with the recording of their responses.

7.2.2 Materials

Visual stimuli and apparatus was the same as Experiment 2, 3 and 4, with the following changes; a five button serial response box was used to record responses. The decision to use a response box was motivated by feedback from the previous study with older adults after they reported that they struggled sometimes to provide an accurate response on a standard keyboard.

Figure 7-1. Example of the three intervening item conditions, Novel, Similar and Repeated intervening items. Under each image is the correct response of new, old and similar to the stimuli going left to right.

7.2.3 Design and procedure

7.2.3.1 Continuous recognition task

Participants completed two tasks, a continuous recognition task, followed by a visual discrimination task. Each trial consisted of one target image pair separated by four intervening items. Target image pairs were either old image pair (an image presented twice in the sequence) or a similar image pair (two different images that share similar visual features). As shown in Figure 7.1, three intervening item conditions were generated; novel (four different images), repeated (one image repeated four times in sequence) and similar intervening items (four different images which were matched according to their categorical label and the degree of perceptual similarity (see Appendix C for additional details on how the similar intervening item group was constructed). In total, participants completed 60 trials, with half of the trials presenting an old target image pair and half presenting a similar target image pair. In total there were 20 trials for each intervening condition. Target image pairs were randomly allocated to one of the intervening item conditions. 10 old and 10 similar target image pairs were presented for each intervening item condition.

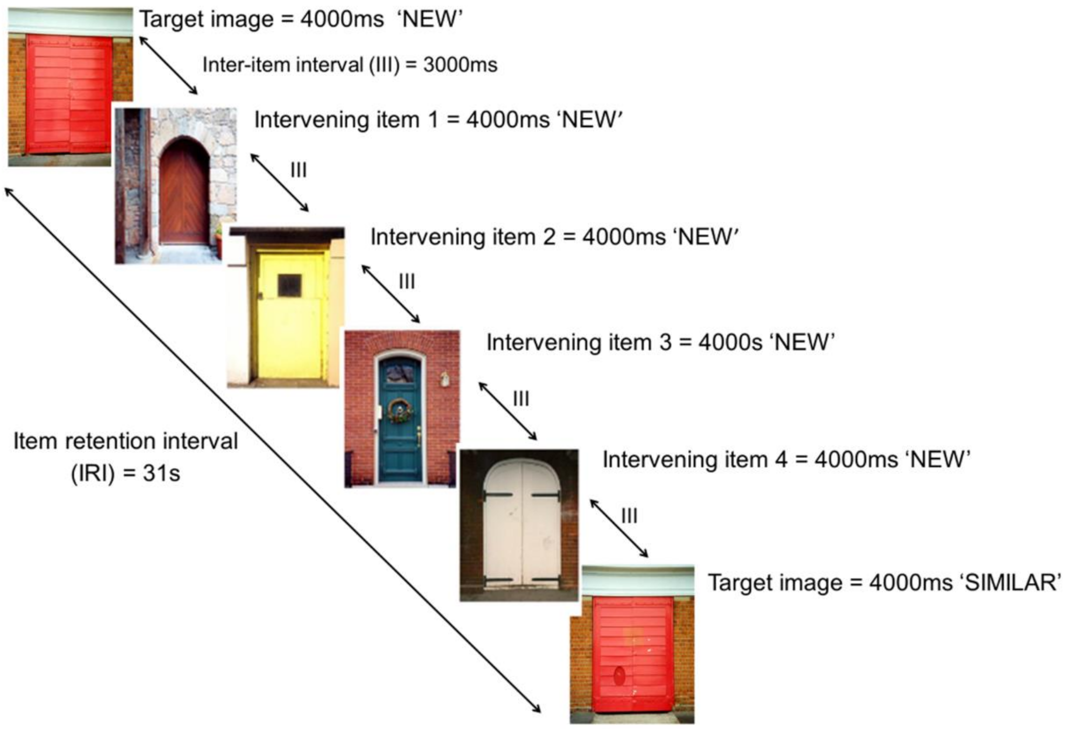

Figure 7-2. Example of the experimental procedure with the novel intervening item group.

For each image presented, participants were instructed to decide if the image was new (first time presented), old (the image was presented previously in the current sequence) or similar (an image that is similar but not identical to an image previously presented in the sequence). As shown in Figure 7.2, each image was presented on screen for a fixed duration of 4000ms. For each image participants were instructed to provide their response using one of three response buttons, clearly marked, 'new', 'old' and 'similar'. Participants were instructed to respond using their index finger on their dominant hand. Each image was separated by an inter-item interval (III) of 3000ms. Taking into account the stimulus duration, number of intervening items and III, each target image pair was separated by an item retention interval (IRI) of 31s. A within subject design was adopted with participants responding to each target image pair (old and similar) and intervening item condition (novel, repeated and similar). Following 18 practice trials, participants completed six experimental blocks of 10 trials. (60 trials in total). Each block lasted approximately 7.05 minutes with a 5 minute break separating each experimental block. Participants completed all the experimental blocks in one session with the distribution and order of the target image pair and intervening item combinations counterbalanced within the session and across participants. On average, the continuous recognition task took 73.75 minutes to complete.

Figure 7-3. Example of the image pairs used in the discrimination task. In bold are the correct responses.

7.2.3.2 Visual discrimination task

Participants were presented with two images, side by side. For each image pair presented, participants had to indicate if the images were the same or different. As shown in Figure 7.3, participants were presented with 90 image pairs consisting of 30 same pairs, 30 different pairs and 30 similar pairs. All the stimuli presented had been seen previously in the continuous recognition task. The same pairs were the old image pairs used in the continuous recognition task and the similar pairs were also unused. The different novel image pairs were constructed using the stimuli from the novel intervening item. Following 20 practice trials, participants completed one block of the discrimination task. Each trial began with a fixation cross of 250ms followed by an image pair. As soon as the image pair appears on screen, participants were instructed to record their response of ‘same’ or ‘different’ using two response buttons, clearly labelled. Participants had as long as they need to view each image pair but were reminded to response as accurately as fast as they could. Once a response was provided the next image pair was presented after a delay of 500ms. The order of the image pairs was randomised across participants.

Table 7-1.
Definitions of the response possibilities for the new, old and similar stimuli; and formulae used for the main derived measures.

Stimulus	Response		
	New	Old	Similar
New	Correct rejection Rate	False alarm rate	Similar bias rate
Old	Miss rate	Hit rate	Incorrect
Similar	Incorrect	Pattern completion rate	Pattern separation rate

Recognition Accuracy Score (RAS) = Z(hit rate) – Z(false alarm rate)
Behavioural pattern separation (BPS) = Z(pattern separation rate) – Z(similar bias rate)
Behavioural pattern completion (BPC) = Z(pattern completion rate) – Z(false alarm rate)

7.3 Results

7.3.1 The key analyses of interfering effects of intervening items

The first analyses (7.3.1 and 7.3.2) ignored age, and then a further analysis (7.3.3) examines age by partitioning participants into younger and older groups. For the continuous recognition task, the mean proportion of responses (new, old and similar) were calculated for each target image and intervening item condition and the three outcome measures calculated, RAS, BPS and BPC (the reminder Table 7.1 shows how these measures are derived). Table 7.2 shows the raw scores, and Table 7.3 the derived measures for the three intervening item conditions (novel, repeated and similar). The interest is in whether one or other intervening stimulus types is most interfering (worsening performance in the derived measures most).

Table 7-2.

Mean proportion of responses (new, old and similar) for each target type (new, old and similar) for each intervening item condition (novel, repeated and similar). Standard deviation is given in parentheses (N = 51).

Target Type		Novel	Repeated	Similar
New	Correct rejection	.76 (.13)	.83 (.13)	.85 (.11)
	False alarm rate	.13 (.12)	.07 (.11)	.09 (.10)
	Similar bias rate	.10 (.09)	.09 (.09)	.07 (.08)
Old	Miss rate	.14 (.15)	.11 (.13)	.06 (.09)
	Hit rate	.77 (.21)	.82 (.19)	.89 (.18)
	Incorrect	.09 (.14)	.06 (.13)	.05 (.12)
Similar	Incorrect	.20 (.17)	.26 (.15)	.35 (.19)
	Pattern completion rate	.56 (.19)	.47 (.22)	.39 (.21)
	Pattern separation rate	.24 (.18)	.27 (.22)	.27 (.19)

For RAS a one-way repeated measures ANOVA with a factor of intervening item (novel, repeated and similar) revealed a main effect of intervening item, F (2, 100) = 40.42, $p < .001$, $\eta_p^2 = .45$. Pairwise comparisons confirmed the RAS was lower for the *novel intervening items* in comparison to the repeated intervening items, $t(50) = -6.70$, $p < .001$, $d = .67$; and the similar intervening items, $t(50) = -8.15$, $p < .001$, $d = 1.02$. The RAS for the similar intervening items was marginally higher than the repeated intervening items, $t(50) = 2.40$, $p = .020$, $d = .26$. In contrast, for BPS, a one-way repeated measures ANOVA with a factor of intervening item (novel, repeated and similar) revealed no main effect, F (2, 100) = 2.19, $p = .117$. However, for BPC the same analysis revealed a significant main effect of intervening item, F (2, 100) = 5.80, $p = .004$, $\eta_p^2 = .104$. Pairwise comparisons showed the BPC for the *novel intervening items* condition did not differ from repeated, t (50) = -.88, $p = .386$; but similar was marginally lower than novel, t (50) = -2.25, $p = .029$, $d = .41$; and similar was marginally lower than repeated, t (50) = -2.82, $p = .007$, $d = .42$. In summary, for RAS, the *novel* condition appeared most interfering; for BPC that role fell to the *similar* condition (these values are highlighted in bold in Table 7.3 for clarity).

Table 7-3.

Derived measures RAS, BPS and BPC for each of the three intervening item conditions (novel, repeated and similar) (N = 51). Standard deviation is given in parentheses. Highlighted values in bold are for statistically significantly lowest performance for the particular measure.

	Novel	Repeated	Similar
Recognition accuracy score	**2.02 (.73)**	2.55 (.83)	2.74 (.69)
Behavioural pattern separation	.62 (.61)	.69 (.71)	.83 (.64)
Behavioural pattern completion	1.38 (.64)	1.46 (.62)	**1.13 (.56)**

7.3.2 Exploring how responding to the intervening four items may have differed – and may account for the key experimental outcomes above

Having addressed the key experimental question – do novel items interfere most – attention may be turned to whether aspects of responding to the intervening items differs – and by implication, whether those differences themselves might underlie or contribute towards any reported differences in interference. First, for each of the four intervening items, in each of the three experimental conditions, a simple calculation is made of whether the response was correct or not (old, new or similar). Table 7.4 shows the simple total correct proportions overall, and Table 7.5 breaks these scores down by serial position, 1-4, and by response type. Of course Table 7.5 shows both the correct response (for example, new for novel) and each of two forms of errors (for example, old for novel and similar for novel); but note that for two conditions, repeated and similar, the identity of this correct response changes following the first presentation (serial position 1).

Table 7-4.
Mean proportion of correct response per intervening item condition. Standard deviation is given in parentheses.

Intervening Item Condition	M (SD)
Novel	.83 (.10)
Repeated	.88 (.15)
Similar	.55 (.13)

For the mean proportion of correct responses, a 3 (intervening item condition: novel, repeated and similar) x 4 (serial position: item 1, item 2, item 3, item 4) repeated measures ANOVA revealed a main effect of intervening item, $F\ (2, 100) = 117.12$, $p < .001$, $\eta_p^2 = .70$; and of serial position, $F\ (3, 150) = 10.11$, $p < .001$, $\eta_p^2 = .18$; and significant interaction, $F\ (6, 300) = 46.65$, $p < .001$, $\eta_p^2 = .48$.

Table 7-5.

Mean proportion of each response (new, old and similar) for each intervening item serial position (item 1, item 2, item 3 and item 4) as a function of the intervening item condition (Novel, Repeated and Similar). Standard deviation is given in parentheses. Correct responses are indicated by (*).

Intervening item condition	Response	item 1	item 2	item 3	item 4
Novel	New	.84 (.11) *	.85 (.11) *	.83 (.13) *	.85 (.10) *
	Old	.06 (.08)	.05 (.08)	.08 (.13)	.06 (.08)
	Similar	.10 (.09)	.10 (.09)	.09 (.07)	.10 (.09)
Repeated	New	.79 (.14) *	.04 (.04)	.00 (.01)	.00 (.01)
	Old	.08 (.11)	.87 (.16) *	.94 (.19) *	.95 (.19) *
	Similar	.13 (.10)	.08 (.14)	.06 (.18)	.05 (.19)
Similar	New	.77 (.13) *	.35 (.18)	.17 (.15)	.19 (.15)
	Old	.10 (.11)	.21 (.12)	.29 (.15)	.29 (.13)
	Similar	.13 (.11)	.44 (.20) *	.54 (.19) *	.52 (.19) *

In each intervening item condition, the correct response for the serial position item 1 was ‘new’. Post hoc analysis of the proportion of correct response for each serial position as a function of the intervening item condition revealed that for the first serial position, item 1, the proportion of correct responses was higher for the novel condition than the repeated condition, t (50) = 3.44, p = .001, d = .39; and higher for the novel than the similar condition, t (50) = 5.12, p < .001, d = .59. For serial position 1, the repeated and similar conditions did not differ, t (50) = 1.59, p = .118. Turning to serial positions 2, 3 and 4, one way repeated measures ANOVAs with a within subjects factor of serial position (item 2, 3 and 4) separately by condition showed no effect of serial position for novel, F (2, 100) = .64, p = .528; but a significant effect for repeated, F (2, 100) = 12.97, $p < .001$, $\eta_p^2 = .21$; and for similar, F (2, 100) = 12.48, $p < .001$, $\eta_p^2 = .20$. Figure 7.4 shows the patterns – a level response for novel, a very modest increase and levelling for repeated, and a marked drop and levelling for similar. The 'take-home' message may be made that responding was at a much lower level in the similar condition across these three items preceding the critical target response in this experiment.

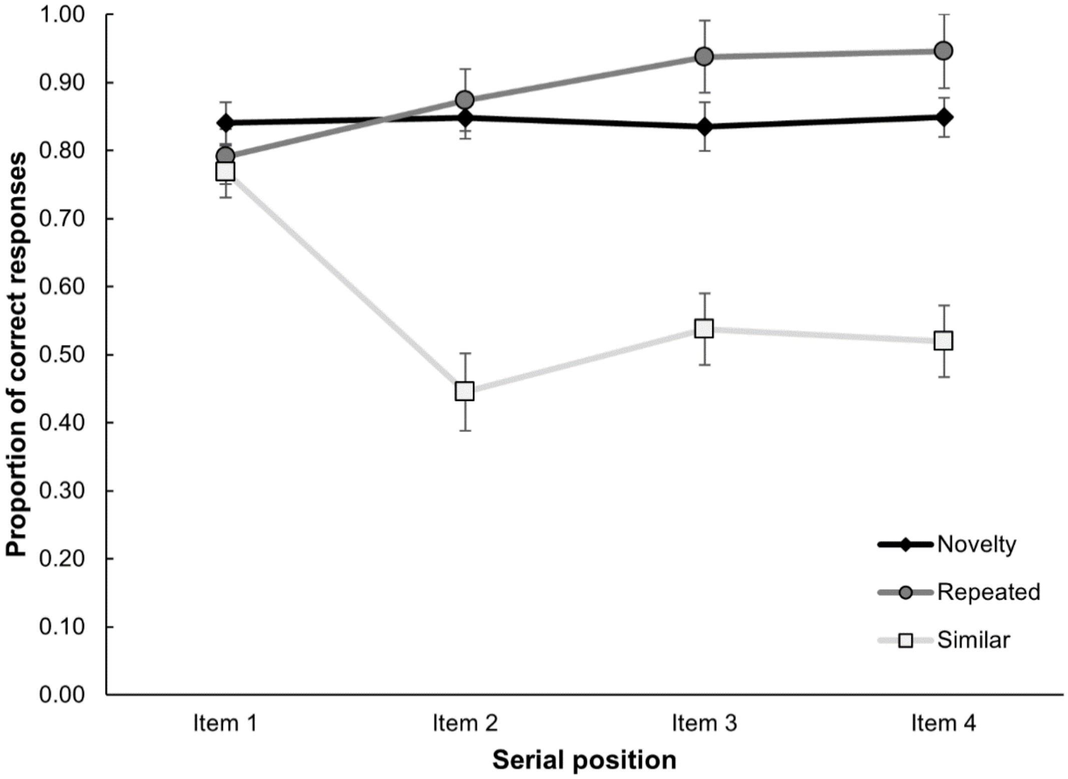

Figure 7-4. Mean proportion of correct responses for each intervening item serial positions (item 1, item 2, item 3 and item 4) as a function of the intervening item condition (Novel, Repeated, Similar). Error bars represent 95% confidence intervals.

7.3.3 Comparison of performance between younger adults and older adults

As explained in Section 7.2.1, participants were partitioned into the younger (N = 28) and older (N = 23) groups. Their raw scores on the task are shown in Table 7.6. A 3 (intervening item condition: novel, repeated and similar) x 2 (age group: younger adult and older adult) repeated measures ANOVA showed no significant interaction between age group and the intervening item conditions for the derived measures, RAS, *F (*2, 98) = .36, *p* = .699; BPS, *F (*2,98) = .10, *p* = .90; or BPC, *F (*2, 98) = .31, *p* = .734. Nevertheless, although the derived measure RAS revealed no difference between age groups (*t* (49) = 1.41, *p* = 1.64), and their underlying hit rates did not differ (younger adults: *M*=0.84, *SD*= .14; older adults: *M*= .82, *SD*S= .21), *t* (49) = .23, *p* = .817; the *false alarm rate* was marginally higher for older adults (*M*=0.13, *SD*= .15) than younger adults (*M*=0.06, *SD*= .05), *t* (49) = 2.29, *p* = .03, *d* = .69.

Table 7-6.

Mean proportion of responses for each stimulus type for the three intervening item conditions (novel, repeat and similar) for younger adults (N = 28) and older adults (N = 23).

		Younger adults			Older adults		
Intervening item condition:		Novel	Repeat	Similar	Novel	Repeat	Similar
New	Correct rejection	.77 (.11)	.84 (.11)	.85 (.09)	.75 (.15)	.82 (.16)	.85 (.13)
	False alarm rate	.11 (.07)	.04 (.06)	.06 (.06)	.17 (.16)	.11 (.14)	.12 (.13)
	Similar bias rate	.12 (.09)	.11 (.09)	.09 (.09)	.08 (.09)	.07 (.09)	.04 (.05)
Old	Miss rate	.13 (.16)	.11 (.13)	.08 (.11)	.15 (.14)	.11 (.14)	.04 (.07)
	Hit rate	.78 (.21)	.83 (.16)	.89 (.15)	.76 (.22)	.81 (.22)	.90 (.21)
	Incorrect	.10 (.13)	.05 (.08)	.03 (.07)	.08 (.15)	.08 (.17)	.07 (.17)
Similar	Incorrect	.18 (.17)	.27 (.17)	.32 (.20)	.23 (.17)	.24 (.12)	.38 (.16)
	Pattern completion rate	.53 (.16)	.40 (.18)	.34 (.19)	.58 (.22)	.56 (.23)	.44 (.21)
	Pattern separation rate	.29 (.17)	.33 (.22)	.33 (.21)	.18 (.17)	.20 (.20)	.19 (.13)

Finally, although one has to step very carefully in considering the raw scores given their inter-dependence (Table 7.6), and analyses in the present thesis have focused predominantly on the derived measures, nevertheless it is of interest to highlight the patterns of responding by younger versus older groups to the demanding *similar* items. Two response types, pattern separation rate (correctly reporting a similar item as similar) and pattern completion rate (incorrectly reporting the similar item as old) are highlighted in Table 7.7, but crucially separately according to the (previously established) stimulus similarity ratings. Without pursuing the patterns of raw scores too far, it may be noteworthy, firstly, that younger adults are performing somewhat better to similar items; but secondly, whilst younger adults were sensitive to the degree of stimulus sensitivity, older adults were less so. As shown in Table 7.7 the younger group performed moderately accurately to similar items, but less so when similarity was high; the younger group performed poorly on similar items, and this did not differ for low or high similarity. These observations were supported by t-test (e.g. correct responses were significantly higher in the less demanding low similarity stimulus pairs than the high similarity pairs in the younger adults, $t(55) = 4.71$, $p < 0.001$, $d = .63$; but this difference was not shown by the older adults, $t(45) = .38$, $p = .704$). Note though that, for both age groups, the distribution of errors to similar items does shift across the similarity manipulation; both make more ‘new’ errors when item similarity is low relative to high. In this sense therefore, the older group are sensitive to changes in stimulus similarity.

Table 7-7.

Mean pattern separation (regarded as target correct responses) and pattern completion rates (regarded as errors) for high and low similar target image pairs across the age groups (younger and older adults). Standard deviation in given in parentheses. Low similarity indicates a more demanding discrimination.

Similarity Rating	Pattern separation rate		Pattern completion rate	
	Younger adults	Older Adults	Younger adults	Older Adults
Low	.36 (.20)	.19 (.18)	.29 (.18)	.39 (.21)
High	.25 (.14)	.18 (.17)	.60 (.16)	.70 (.19)

7.4 Discussion

The aim of the this final experiment was to test the idea that successive item *novelty* maps onto greater inter-item interference in the task. In confirmation, performance for both old and similar items was affected by the intervening items' novelty manipulation. Firstly, the ability to accurately recognise an old target item declined when the intervening items were four unique novel items, in comparison to both the repetition of a single intervening item as well as four items which share similar visual features. When separated by similar intervening items, performance was better in comparison to the novel and repeated intervening item conditions, in terms of both an increase in the accurate identification of old items and a decline in the number of similar target items falsely recognised. These outcomes are the first to show that the perceived novelty of each item encoded during a continuous recognition task may determine the degree of interference generated – and in turn performance on the task – but see the moderating comments below. A secondary aim was to test for an interaction between the presumed novelty-based interference and age. There was no interaction between the intervening item conditions and the age groups, although overall performance was somewhat better in younger adults than older adults, replicating the findings of the earlier experiments. Can we therefore conclude that item novelty is itself introducing a confound in the BPS task? Possibly. Unfortunately, as explored in Results above, the present experimental design was flawed. Not only did the design manipulate the stimulus types preceding the critical target item, it also inadvertently introduced differences in response levels across those preceding items. This is unfortunate for our confidence in the novelty manipulation.

Finally, a further analysis, following up the interest in age and responding to similar items in Experiment 5, took advantage of the fact that the participants in Experiment 6 spanned a wide age range. By partitioning the sample into younger and older age groups, it was possible to further explore similar response patterns in each group (the usual cautionary reminder is made in dealing with selected raw scores in the three-response task). This exploration actually revealed an attraction of the three-response task over the new/old two-response task used in Experiment 5. Notably, the pattern of errors in Experiment 6 to *similar items* could take two forms, rather than just one: reporting similar items as old (which is of course 'pattern completion rate') and reporting similar items as new (termed simply 'incorrect'). Now, when stimulus similarity was studied, older participants' correct similar responses were at a very low level for both low and high similarity stimuli; only the pattern of errors revealed that, indeed (unlike the outcome for the new/old task in Experiment 5) the older participants here were, like the younger group, indeed sensitive to stimulus similarity – they made more 'new' response for the items of low stimulus similarity.

Chapter 8 General Discussion

The hippocampus lies in the medial temporal lobes and is perhaps the most studied neural structure within the brain within the psychological neurosciences. Over decades functional descriptions of hippocampus have come from many levels of scientific enquiry, with a fairly recent text attempting integration across neural, anatomical, cognitive and physiological data sources (Andersen, Morris, Amaral, Bliss, & O'Keefe, 2007). Even a casual reading of this text would reveal that the hippocampus is involved in memory processes (although its role in specifically spatial memory processes appears to be predominant in some descriptions). The present very narrow focus nevertheless embraces two big problems facing us if we are to have a better understanding of such memory processes – what is the inter-dependence of encoding and retrieval? As discussed in this final chapter, this necessary inter-dependence presents significant difficulties of interpretation of response patterns (and response criterion biases) in the three-response task, and recommendations will be made of one solution to this.

In trying to answer the question of the role of hippocampus in immediate memory more broadly, evidence has converged from many sources - behavioural, anatomical, neurophysiological and computational modelling - that specialised circuits support distinct encoding and retrieval via the processes of pattern separation and pattern completion (Hunsaker & Kesner, 2013). The present work has accepted the general view that pattern completion is the retrieval of a memory for an item based on a degraded or partial stimulus cue (Hunsaker & Kesner, 2013). Such a memory process may rely on the auto-associative properties of hippocampus (notably the recurrent collaterals of region CA3) to achieve pattern retrieval based on partial activations across a network (e.g. Rolls, 2007). Pattern separation in contrast, is the process whereby similar or overlapping representations are stored as distinct orthogonal memories, reducing the overlap and interference from previously stored items (Yassa & Stark, 2011). It is believed to occur in the hippocampal dentate gyrus and CA3 region (Deuker et al., 2014). The thesis has addressed a set of six questions of the application of the three-response BPS task in revealing aspects of these fundamental processes as they are manifest in discrimination and generalisation across sets of visual stimuli.

8.1 Summary of the experiments

Two areas of research have been called upon in the present attempt to grapple with this question of better understanding these processes of encoding and retrieval. One is an exciting recent and growing set of demonstrations that certain circuits governing immediate memory encoding and immediate memory retrieval may be tapped using a continuous recognition task with human participants having three simple responses (and the task has the attraction that the data is relatively easy to collect). The other literature is contemporary memory theory as it pertains specifically to immediate memory forgetting. No attempt is made here either to review these literatures – they are enormous – but rather, an attempt has been made to identify some key relationships (such as the role of short-term memory trace consolidation operating across several seconds, or the role of inter-item interference) that allow us to better grasp what the contribution of the three-response task is – and also, hopefully, to provide some recommendations to improve the task, or at least to provide researchers with some constraints that they would do well to satisfy. These goals have been achieved to some extent.

The design of each study will now be briefly described and their results summarised. In Experiment 1 was a direct test of claims that clinical populations (normal, aMCI, AD) were distinguishable by a manipulation of a variable termed ‘lag’ in the BPS (number of intervening items separating ‘old’ and repeated items or similar items). In Experiment 1 a key interest was to manipulate lag whilst maintaining a fixed memory retention interval, a critical variable in contemporary accounts of forgetting termed ‘decay interval’. This was achieved by manipulating the time interval between successive items (recall the conditions were labelled 4_short, 4_long and 12_short where 4 refers to number of intervening items and short/long to memory retention interval). Performance in a sample of young adults on the task was surprisingly high – and did not vary appreciably across experimental conditions varying number of intervening items. It was supposed in discussion of that experiment that this high performance, and resistance to disruption by potential interferers, might be attributed to the availability of verbal labels for the everyday objects, permitting a verbal code maintenance strategy. Derived measures of performance in the task have been rehearsed many times in the thesis, with reminder tables throughout: RAS, BPS and BPC. Three features of the experiment were identified in statistically marginal performance comparisons on these key measures: interference through increasing number of intervening items from 4 to 12 was absent; forgetting through increasing decay interval from 20.5s to 56.5s was absent; however, a significant positive effect of post-item interval was observed.

In the same design in Experiment 2, using stimuli belonging to a single object category (so limiting a verbal rehearsal strategy), task performance was determined by the number of intervening items as well as the post-item interval. In this new version of the task, both measures RAS and BPS were low overall, and even declined slightly as the number of intervening items was increased. When the levels of interference was arranged to be equal between tested item pairs, performance improved with an increased inter-item interval. However, one factor was noted - as hits increased (identifying old items correctly), so too did false alarms (reporting similar items as old).

In Experiment 3, a group of older adults (N=22, M_{age}= 77.74years) completed the three response task with the new same-category stimuli. For this group performance (as indexed by their ability to accurately identify both old and similar stimuli) modestly improved as the inter-item spacing between stimuli was increased. To explore the age variable, this group were compared to the younger sample from Experiment 2 on the same stimuli and experimental design. Performance was somewhat better in the younger sample for RAS, BPS, and BPC in separate statistical tests of 4_short, 4_long and 12_short (with just the occasional moderating word ‘marginal’ in the t-tests). However, an inspection of the raw scores underlying the derived measures indicated that the older sample appeared insensitive to stimulus novelty: errors to ‘new’ items were higher than the younger sample, both in their errors of reporting new items as similar (similar bias rate) and of reporting new as old (false alarms).

In Experiment 4, the benefit of the inter-item interval on task performance was examined further by manipulating the opportunity to engage in short-term memory trace consolidation. A new sample of younger adults completed a novel variation of the three response task; here a secondary task was introduced during the interval between each item. By varying the onset of this secondary task, the time available to engage in short-term memory trace consolidation was manipulated. An extended consolidation window led to improved performance in this group as indexed by their ability to accurately identify old and new stimuli, as shown in the RAS. However, this apparent benefit of short-term trace consolidation was not apparent in responses to similar stimuli (benefits of extending the post-item interval were not observed in the BPS and BPC measures).

In Experiment 5, attention turned to the important manipulation of perceptual similarity on responses to similar items. The idea this term is intended to convey is that the number of physical features shared by the stimuli in the task may define the percept along a continuum from ‘not at all similar’ to ‘highly similar’. To determine the perceptual similarity amongst the new door items, an independent sample of younger adults rated sets of similar pairs in terms of their perceptual similarity on a 7-point scale. Next a new younger sample completed a two-response version of BPS (Experiment 5b). Here stimuli varied on the three dimensions (new, old, similar, the latter according to the 7-point scale) but the responses were a binary ‘new’ and ‘old’. It is easiest to repeat the summary of this data from Chapter 6, that the outcomes were promising, if not fully realised. So performance in the new/old task did approximately map onto the 7-point similarity scale (but note the ‘marginal’ qualification in the t-tests). Subsequently, Experiment 5c was a re-analysis of selected already reported outcomes of the three-response task in light of the new assignment of the door stimuli according to mnemonic similarity, using both data for a younger sample (Experiment 2) and an older sample (Experiment 3). In this re-analysis, the focus was only on accuracy (pattern separation rate: accuracy of reporting similar to similar items) and errors (pattern completion rate: errors in reporting old to similar items). For accuracy, performance did map onto the similarity scale appropriately (but again, note the ‘marginal’ in the t-tests). For errors there was a statistically reliable and important contrast: whereas for the younger group errors were significantly higher as item similarity increased (thus they were showing sensitivity to feature overlap), for the older group the low and high similarity items were not distinguishable.

In Experiment 6, the role of inter-item interference on task performance was examined through the lens of stimulus novelty. The idea behind this was that highly novel events are able to capture attention and demand processing time and resources (see a description of one recent account of how this occurs below) so that they are especially disruptive with ongoing memory trace maintenance (again, see below for why this may be). It was arranged that target item pairs (old and similar) were separated by four intervening items which varied in their degree of novelty. The outcomes were interesting. Firstly, the ability to accurately recognise an old target item declined most when the intervening items were four unique or novel items. In contrast similar intervening items appeared to offer less interference. However, conclusions were moderated by the recognition that the design of the experiment was flawed – the performance levels differed across experimental conditions to those intervening stimulus items. Finally, the participants in Experiment 6 were partitioned by age and, following up the observations in Experiment 5, responses to similar stimulus items were examined according to age and according to whether stimuli were (by the ratings' data) judged of low similarity or high similarity. Two noted findings were shown for accuracy (similar responses to similar items): firstly, whereas the younger group showed modest but reliable performance in accuracy, the older group were effectively at chance level; secondly, the younger group reliably improved their performance when stimulus items were more dissimilar, but the older group did not differ across stimulus similarity. The findings for errors were rather similar for younger and older groups: errors of reporting highly similar items as 'old' shifted to reports of 'new' when the items were of low similarity.

8.2 Some observations on pattern separation and completion and the 'success' of the experimental manipulations

As noted, the theoretical framework supporting pattern completion (retrieval) and pattern separation (encoding) relies on diverse experimental methodologies, and recent reviews (Hunsaker & Kesner, 2013) broadly support the current operational descriptions of the processes, and their realization within hippocampal circuits. The human behavioural index of these forms of memory encoding and retrieval in human studies has been the Behavioural Pattern Separation Task (BPS) (Kirwan & Stark, 2007). In the standard task that guided the present work, stimuli are pictures of everyday objects (a wheelbarrow, an apple, a table), which are presented one by one in sequence on a computer screen and observers report for each item whether it is novel (new), previously viewed in the sequence (old), or whether it shares features with a previously viewed item (similar). The index of "behavioural pattern separation" is then calculated as the difference between the rates of ‘similar’ responses to similar items minus 'similar' responses to novel items. Pattern completion is, strictly, recall of an item following presentation of a partial cue, but within the standard BPS task is taken to be captured by the rate of similar items incorrectly identified as 'old' (Toner et al., 2009). The reasoning here is that when a similar item is identified as old, it has acted as a partial cue for the recall of a prior item.

Although the present thesis has maintained the early terminology in the literature of these two terms, pattern separation and pattern completion, for the outcomes of a purely behavioural task, no attempt is made to force too close a correspondence between discriminative abilities in the three-response (or two-response) task and the hippocampal circuits or memory processes where those terms properly apply. Indeed, some of the original investigators of the behavioural ‘assay’ on those memory processes have themselves moved away from the terms and now prefer to emphasise the discriminative element and suggest ‘mnemonic similarity task’. This certainly has the attraction of making clear that only loose parallels in memory mechanisms or processes are being suggested, but unfortunately there are very many behavioural memory discrimination tasks that might fully fall within the breadth of this new suggested term, so the present author has preferred to adhere to the original labels. In writing about the discriminative performance, in interpreting data, however, it is always borne in mind that it is a big jump from discrimination to neural circuitry.

Another ‘jump’ is making inferences about the effects of ageing in the performance on the BPS. Several studies have explored the link. Thus, one aspect of the greatly renewed recent interest in pattern completion and pattern separation arises from attempts to understand older adults' memory and their forgetting. Indeed, a growing body of evidence suggests that older adults exhibit a deficit in encoding new memories so that they are distinct from previously stored items. In other words, they show impaired pattern separation (Carr et al., 2015; Holden et al., 2013). Indeed, recent research suggests that visual recognition impairments in older adults are due to an impaired ability to identify stimulus novelty (Yassa & Stark, 2008). As we age, ‘false’ or impaired recognition of everyday objects increases (Norman & Schacter, 1997), which can result from novel items being viewed as though they had been previously seen (Koutstaal & Schacter, 1997). Yeung et al. (2013) presented older adults with a series of everyday objects in an initial study phase. They then recorded their eye tracking behaviour whilst they viewed some of the objects from the study phase amongst new objects, which shared either high or low similarity with previously viewed objects. Commonly, mean eye fixation is greater for the exploration of a novel object (Henderson & Hollingworth, 2003). Yet Yeung et al. found that mean eye fixations in their older adults did not differ across old repeated items and items sharing high similarity - apparently new items were falsely viewed as old. This may have reflected either impoverished encoding in the study phase (Molitor et al., 2014) or failure to identify novel visual features.

Toner et al. (2009) used the BPS to compare pattern separation performance between younger adults and healthy older adults, aged over 65. Groups did not differ in their performance for 'old' and 'new' items; but for 'similar' items the older adults performed more poorly. Toner et al. argued that age related changes to the hippocampus in older adults may result in *inefficient pattern separation*, rather than a recognition memory deficit per se. Stark et al., 2013) demonstrated that behavioural pattern separation scores in the BPS gradually decline across the lifespan. Based on this, it could be suggested that diminished ability to recognize the novel features of objects results from neurocognitive ageing, (i.e. Wilson et al., 2006), where in older adults the formation of new memories is hindered by interference from prior memories. This question of age is explored further in this chapter, but for now it is noted that the findings are a little ambiguous. Whereas the new/old task data in Experiment 5 reveals no differences in errors for similar items in the older group across low and high similarity stimuli, the new/old/similar task data in Experiment 6 does reveal a difference: like the younger group, the older group errors shifted from reporting similar items as ‘old’ when stimulus similarity was high to ‘new’ when stimulus similarity was low.

Besides the terms adopted, and caution in making inferences about neurophysiological processes underlying the encoding and retrieval of recent material, the present thesis has emphasised that caution needs to also be paid to the application of the task to clinical distinctions (such as between aMCI and AD), to the validity of the measures, and perhaps most particularly to the role of forgetting in immediate visual memory that is known to be both time-based and interference based (e.g. Ricker et al., 2016; McKeown & Mercer, 2012). Introduction provided a brief tour of some of the most important or influential contemporary proposals for short-term forgetting. The key ones were time-based trace decay (perhaps the oldest of the suggestions in the forgetting literature, and still hotly debated (cf. Ricker et al., 2016), release from stimulus interference brought about by greater temporal spacing of stimulus items (one version of this proposal being the temporal distinctive model), and a form of memory trace consolidation termed here ‘short-term consolidation’ (since, as Wixted, 2004 discusses, the broadly used term ‘consolidation’ may have different time courses). In light of each of these forgetting mechanisms, the indexes of ‘separation’ and ‘completion’ may be better understood. Certainly, short-term consolidation benefits encoding of old items in the BPS task; but it is undoubtedly benefiting performance for ‘similar’ and ‘new’ items too. Similarly, inter-item interference will moderate accurate encoding, but it will have implications for pattern completion too: as shown in the final experiment in the present series, ‘hits’ (recognition of an item presented in the recent past) may be moderated by how novel the immediately preceding items are. It was tentatively proposed in Chapter 7 that novelty was engaging of attention and hence, novel items in the continuous recognition task may be more interfering with the short-term memory trace of earlier items. It was acknowledged, however, that design flaws (notably, the differences in response levels to the intervening items in Experiment 6) limited confidence in this idea. Nevertheless, the third mechanism of forgetting, inter-item spacing, manipulated in Experiment 1 and 2, offers more confidence in the conclusion reached – the benefit of extending the interval following a stimulus item was marked. This conclusion was even strengthened by the observation that it manifested itself *despite* the overall memory retention interval being extended, inviting trace decay (if this process is real). This last conclusion leads to the summary conclusion, that it may be premature to assign performance profiles across differing clinical populations to statements about ‘reduced pattern separation’ – the populations, simply, may differ in their ability to encode items given only a second or so between them. Others, interestingly, have recently acknowledge that, whilst potentially fruitful, caution needs to be exercised in linking the BPS to clinical conditions or assignments in terms of memory impairments (Leal and Yassa, 2018). These authors raise the perhaps obvious point that numerous other neurobiological processes besides pattern separation may contribute to discriminative performance.

8.3 The problems of verbal rehearsal and longer-term associations

Turning to the present considerations of achieving a more valid form of BPS, one key concern was the opportunity for verbal categorisation or labelling in the standard BPS using pictures of everyday objects. The use of verbal labels can be an effective strategy to improve visual recognition memory, as categorising memoranda forms a less noisy representation (Lupyan, 2008). For mixed-category visual stimuli, verbal labels can facilitate memory recall (Lupyan, 2008; Richler et al., 2013) and over short time intervals, protect the memory trace from decay and interference through active rehearsal (Berman et al., 2009). When investigating visual memory often verbal rehearsal is eliminated by introducing articulatory suppression during each inter-item interval (Lewandowsky & Oberauer, 2015). However, Sense, Morey, Prince, Heathcote and Morey (2017) highlight that for visual discrimination experiments there may be no need to introduce an ‘articulatory suppression’ control at all; they demonstrate that for their simple change-detection task with coloured squares no verbal strategy is likely or indeed exploited by participants. Certainly, it is reassuring that feedback from the participants in the present experiments with the new door stimuli both after the practice trials and following the study, supported our view that they did not employ a verbal strategy to perform the task (none reported adopting a strategy such as ‘a blue church door’). The more important observation is, in agreement with the argument of Liu et al. (2016) for achieving valid stimuli in the BPS, that the use of everyday objects is likely to cue retrieval of past associations, engaging longer-term memory and therefore confounding the enterprise. These authors recommend the use of never-experienced-before abstract stimuli. The present stimuli do not quite achieve this. They are not abstract and, whilst a particular instance of a door may never have been experienced before, nevertheless the category is all too familiar. What is reasoned here, however, is that task strategies or discriminative performance will not arise as a result of the general class of ‘door associations’ and the stimuli will not permit of successful verbal rehearsal. It is, though, acknowledged that some form of quite abstract images might offer even more valid stimuli (McKeown et al., (2014) have indeed designed a promising stimulus set based on images undergoing a filtering transform to produce unique abstract ‘modern art’ images). Yet these more abstract images might introduce another difficulty – achieving parametric alteration of the degree of interference across stimuli (like the similarity index reported in Chapter 6).

8.4 Some words of caution in interpretation

Consider first the underlying processes that the BPS aims to reflect. The theoretical developments of the role of the hippocampal system in episodic memory which are progressing rapidly (Leal & Yassa, 2018 report an exponential increase in publications on pattern separation since the mid-seventies) arise within the neurosciences from, for example, studies of the effects of lesions to hippocampal circuits, of fMRI recordings during discrimination tasks, and of animal models. One aspect is the influential claim (Rolls, 2013) that a form of auto associative network is realized by hippocampal CA3 neurons to permit pattern completion based on partial cues. In artificial neural networks, ‘error’ in output is fed back to adjust weights within the network until it ‘finds’ the correct output based on incomplete input – in other words the network auto-associates the output from the input given. In addition it is claimed the sparse connectivity of mossy fibres to CA3 has a randomizing effect on the representations within that network so that the patterns stored are as different from one another and from other patterns as is possible to allow pattern separation. Here the idea is that a small pool of features may be spread across a broad field so that the chances of any two feature elements overlapping is very much reduced. This spreading out of patterns within CA3 will mean that interference between similar pattern memories is minimized.

Current understanding of the proposed hippocampal pattern completion and pattern separation processes based on human behavioural discrimination learning and retrieval may be an exciting development – but some caution is needed in interpretation (Hunsaker & Kesner, 2013). This thesis (Chapter 2) has reported that the comparisons of older adults, individuals with aMCI and with AD in the study of Ally et al. (2013) appeared to uncover a form of rapid interference-based forgetting, with individuals with aMCI and AD showing decreased pattern separation scores (BPS) with increasing 'lag'. However, as noted, Ally et al’s manipulation of lag confounded with some very basic factors well-known to promote forgetting in the short term, such as time-based decay. Here too the role of short-term memory consolidation (Wixted, 2004) was highlighted (and confirmed in Experiment 4). Such a conclusion is fully in line with the recent observation by Molitor et al. (2014), using an eye-tracking measure to assess false alarms in a continuous recognition task like that used here: apparently errors in the task of reporting 'old' to similar items were associated with decreased fixations during the earlier presentations of those items. In a temporally densely packed sequence of items, the problem presumably was one of *insufficient encoding of items into memory*. A natural assumption therefore might be that one form of impairment within aMCI and early AD is encoding or consolidation time for items within a continuous recognition task, rather than failures in pattern completion and pattern separation per se. The next section explores this idea further.

In fact recent evidence points to a variable termed ‘rate of learning’ as a key feature of aMCI and early onset dementia (Hanseeuw et al., 2011; Walsh et al., 2014); whilst Wang, Li, Li and Zhang (2013) have identified the key variable under conditions using massed trials in aMCI as the impairment in encoding into memory of successive items. Alternatively, it may simply be that memory precision gets worse with age and with onset of mild cognitive impairment. Studies in the working memory literature using continuous reproduction tasks have shown that even to retain a single item in memory (e.g., a single colour patch), older adults show lower precision (Peich et al., 2013; Pertzov et al., 2015; Souza, 2016; Zokaei, Burnett Heyes, Gorgoraptis, Budhdeo & Husain, 2015). Therefore, it is possible to view the BPS as a measure of memory precision of an item (how precisely do I remember the item I saw before) so as to be able to tell that a subsequent similar item is “just” similar and not exactly the same. In this sense it is relevant that the data of Toner et al (2009) show that older adults reported similar objects more often “old” rather than “similar”, whereas younger adults showed the reverse pattern. This is in line with the idea that the older adults showed lower precision for the stored information within memory.

So caution needs to be exercised for many aspects of BPS. If it is adopted as a tool to categorise clinical populations, what is it measuring and, of these populations what differences in susceptibility to within-item interference, item trace decay, item short-term trace consolidation, or indeed variations in ongoing attentiveness may govern discriminative outcomes? As one of a toolbox of measures, and armed with normative BPS data, the clinical utility may be highly promising of course. However, as argued throughout this thesis, the real promise will only be realised if task and stimulus-set validity are achieved, such as using sufficient inter-item temporal spacings (and maybe a limited number of stimuli per session), avoiding verbal rehearsal through single-category or abstract stimuli, parametrically varying inter-item interference, and perhaps too taking note of the identity of the item (new, old, similar) immediately preceding the target item. Yet, however many words of caution, it is reassuring that, unlike common discrimination yes/no memory tasks, the introduction of the ‘similar’ does open exciting opportunities for future investigators to better understand the ‘mnemonic rigidity' on the one hand and 'mnemonic flexibility' on the other of differing populations (Leal & Yassa, 2018).

8.5 Don't forget about forgetting

One observation in the gestation of this thesis was that, ironically for an area of study into memory, the literature on BPS was curiously silent on the possible role of forgetting in governing participants' discriminative ability. Some of the recent evidence demonstrating that non-verbal short term memory declines as a function of time passing between encoding and recall was outlined in Introduction (e.g. McKeown et al., 2014; Mercer & McKeown, 2014; Ricker & Cowan, 2010, 2014; Zhang & Luck, 2009). Also mentioned there was that the precise role of the elapsing time is still debated within the literature (Altman & Schunn, 2012; Mercer, 2014). One suggestion is that time protects memory traces from ongoing interference by isolating stored items in memory. For instance, it can be supposed that each item presented in the continuous recognition sequence in BPS represents an event along the continuum of time (Shipstead & Engles, 2013). For the accurate recognition of old and similar items, two separate time intervals are in operation, namely, the retention interval between encoding and recognition, and the interval separating each item. As already noted, one account of the relationship between these two intervals is the theory of temporal distinctiveness (Brown et al., 2007; Ecker et al., 2015) which emphasises that within a fixed retention interval, increasing the temporal isolation of each item in a sequence will reduce confusability and manage the proactive interference generated from previously stored memory items.

Broadly, temporal distinctiveness theories (e.g. Bjork & Whitten, 1974; Brown et al., 2007; Burgess & Hitch, 1999) emphasize the role of *relative time* (Mercer, 2015) suggesting that the distinctiveness of a memory item is determined by the distance separating it from other items along a temporal dimension (Brown & Lewandowsky, 2010). However, as time elapses, the temporal distance of each item becomes logarithmically compressed (Grange & Cross, 2015), reducing the distinctiveness between items as they recede into the past (Lewandowsky, Brown, Wright, & Nimmo, 2006). Successful recall is therefore determined by both the retention interval and the temporal isolation of each item on successive trials (Cowan, Saults, & Nugent, 1997; Unsworth, Heitz, & Parks, 2008). Recent interpretation of temporal distinctiveness, (e.g. Brown et al., 2007) emphasises a ratio rule (Glenberg, Bradley, Kraus, & Renzalia, 1983), proposing that successful recall of a given item is determined by the ratio between the inter-item interval between each item and the retention interval between initial encoding and retrieval (Ecker, Tay et al., 2015; Oberauer & Lewandowsky, 2008; Souza & Oberauer, 2014). As this ratio increases, memory items become less distinct, resulting in poorer memory performance.

There is emerging evidence supporting temporal distinctiveness models when applied to visual memory performance (e.g. Guérard, Neath, Surprenant, & Tremblay, 2010; Shipstead & Engles, 2013; Souza & Oberauer, 2014). In a recent visual array task, Shipstead and Engles (2013) demonstrated that participants had difficulty detecting changes to a studied memory array (four coloured squares) at longer retention intervals, especially when there was only a short inter-trial interval separating the current memory array and the previous trial. These findings were extended by Souza and Oberauer (2014) who varied the distinctiveness of memory items by carefully manipulating the ratio between the retention interval and the inter-trial interval. Using a colour recall task, the interval between a memory array of six coloured circles and a single test cue for colour judgement was separated by either 1000ms (short) or 3000ms (long). Each trial was separated by an inter-trial interval of 1000ms (short) or 7500ms (long). The probability of correctly recalling the cued colour was equivalent for similar distinctiveness ratios (i.e. short: short and long: long). Performance was best for a relatively short retention interval and a long inter-trial interval, rendering each memory array temporally distinct. Note that the time-scale of this study is really very close (for stimulus presentation time and inter-stimulus spacing) to BPS.

Actually, there is data inconsistent with the ratio rule here in the laboratory at Leeds (McKeown et al., 2014; see also Mercer & Duffy, 2014). In McKeown et al. (2014) participants were presented with a target array of two abstract shapes rather like small coloured abstract paintings. Following a retention interval, participants were asked to decide whether a single test picture presented was the same or different to the target array. This test picture was either a positive probe (same as one in the target array), a negative probe (previously unseen item) or a recent negative probe (an item that had occurred within a target array on a prior trial). The recent probes task has a single highly attractive feature – namely it permits experimenters to test the strength of an enduring memory trace from earlier trials whilst making highly unlikely that participants have engaged in any form of active rehearsal or attentional maintenance strategy. Thus, if we observe slowed responses on the recent-negative probe trials, we may suppose the memory trace of that item from an earlier trial is actively inhibiting the ‘no’ response on the current trial. Similar to BPS, the recent negative probe may introduce proactive interference, causing slower reaction times in contrast to the negative probe. Across two experiments, McKeown et al., reported such slowed recent negative responses even where the interval separating trials was quite extended – more than 6 sec (in fact if we calculate the total interval from the item on the prior trial to the onset of the test probe, the memory interval is just over 20 sec!). In summary, sometimes a memory trace of a visual item resists loss of fidelity over many seconds even in a sequence of other items, and sometimes (e.g. Souza & Oberaurer, 2014) it is vulnerable to interference.

It is worth mentioning finally in this Section that a difficulty in ‘applying’ models such as the conventional temporal distinctiveness ratio rule to BPS. If we consider the relative intervals – the inter-trial interval relative to the current trial retention interval – there are certain situations where that rule cannot be applied in a straightforward manner. The most obvious case is where the memory retention interval contains a series of additional memoranda (like BPS). As such, the delay between item presentation and recall is filled with multiple other items - and so predictions of the ratio rule are made confusing. It is of course possible for us to calculate a similar distinctiveness ratio, where the retention interval is between the first presentation of an item and its subsequent ‘repeat’ or ‘similar’ item probe; and the time elapsed since the previous trial is simply our inter-item interval. However, where ‘the memory judgement’ is compared with items earlier in the sequence such a straightforward ratio of time intervals is probably over-simplistic as it ignores intervening items. So, despite the growing body of evidence in support of the temporal distinctiveness account (Ecker, Brown, et al., 2015; Ecker, Tay, et al., 2015), the present author is still unsure how it would predict the current findings.

8.6 Time for Consolidation?

There is a well-rehearsed idea that 'time for consolidation' enhances short-term memory (Bayliss et al., 2015; Jolicœur & Dell'Acqua, 1998). During an unfilled time interval, memory consolidation can be defined as an active process that works to strengthen a new memory trace so that it can be successfully retrieved at a later point in time (Dewar et al., 2014; Mercer, 2015). Arguably, visual memoranda suffer from rapid time based decay if there is reduced opportunity for engaging in a consolidation process (Knöchel et al., 2015); as a result, the memory trace is more vulnerable to interference or overwriting from succeeding items (Nieuwenstein & Wyble, 2014, offer a recent test of so-called short-term consolidation). One influential champion of memory consolidation has been Wixted (2004). As this author points out, the concept is hardly new and may be traced to a forgetting law of Jost at the end of the 19th century captured by the insight that, with elapsing time, old encoded items within memory become less vulnerable to the disruptive effects of subsequent events: in other words, they show an ever-slowing proportional memory decay function. This temporal gradient of retroactive interference, whereby allowing a temporal interval free of interfering material post encoding strengthens the memory trace, appears a reasonable one intuitively – but it also has support experimentally. Thus, using visual characters (unfamiliar written items) as memoranda and varying time between items, Ricker & Cowan (2014) observed that limiting post-encoding consolidation time using brief inter-stimulus intervals, impaired memory recall. They concluded that “whether or not time-based forgetting will be observed in a working memory task is largely determined by the amount of time allowed for consolidation of working memory” (p. 427). Similarly, Bayliss et al., (2015) varied post-encoding time for lists of consonants by introducing a demanding processing activity either immediately or following a delay (whilst equating retention interval across conditions); they interpreted the impaired performance in the immediate condition as consistent with a consolidation process. Unfortunately, as a careful reading of a recent review (Ricker, 2015) of consolidation in short-term memory makes clear, there is a surprising lack of clarity as to the time-course of short-term consolidation or indeed whether and how it might differ from the more familiar 'encoding time' of the memory trace. While a solely interference-based account may be insufficient for the entire pattern of the present findings, further research is urgently needed to develop a clearly defined account of consolidation in visual short term memory.

The first proposal of a consolidation process operating in memory occurred more than a century ago. Muller & Pilzecker (1990) noted that the learning of new information can disrupt learned information, provided that this follows on in quite close temporal proximity, ensuring that the ‘memory’ items are in a relatively fragile state. (McGaugh, 2000). Across neuroscience and psychology, consolidation has been subject to a considerable research effort, looking at the loss of information over seconds, hours, days and even years (e.g. Dudai, 2004; Dewar et al., 2014; McGaugh, 2000; Stickgold & Waker, 2005; Wixted, 2004). Given these ranges, any examination of consolidation must clearly define the context and time scale of this process.

When presented with a stimulus, the process of encoding generates the initial memory trace, establishing the identity and features of the stimulus (Ricker, 2015). Encoding persists so long as the focus of attention is directed towards the memory trace or the presented stimulus (Woodman & Vogel, 2008). The encoding process terminates when the trace is generated in short term memory or in the absence of the physical stimulus (Ricker, 2015). It is argued that in this initial state, the memory trace is vulnerable to rapid loss unless it undergoes further strengthening via short term consolidation (Ricker & Cowan, 2014). Short term consolidation relies on directed attention towards the memory trace and it occurs in the absence of bottom up support from the physical stimulus (Jolicœur & Dell’Acqua, 1998). Some claim that ‘short-term’ consolidation is complete within about two seconds, although this relies on some particular assumptions (see Nieuwenstein & Wyble, 2014). This is distinct from long term consolidation, a slower process of stabilisation of a memory trace into long term storage. This process is observed over hours and days, via the mechanisms operating between the hippocampus and cortex (Eichenbaum, 2000; Remondes & Schuman, 2004). Presently the literature appears confused as to the temporal properties of short-term consolidation, and has not reached consensus as to its properties, or indeed how it may be distinguished from maintenance processes such as rehearsal or refreshing.

Recently, the role of consolidation has been the subject of study within in short term memory (Bayliss et al., 2015; Dewer et al., 2007; Nieuwenstein & Wyble, 2014; Ricker & Cowan, 2014). Short term consolidation by these authors appears to be the process that strengthens a given memory trace (Ricker, 2015), to enable both accurate retention and retrieval at a later time point (Dewar et al., 2014; Mercer, 2015) – although what ‘strength’ is, is left unsaid. Following the identification and encoding of a visual stimulus, some authors have relied on the idea that the resulting memory trace is subject to time based decay (Bayliss et al., 2015; Ricker & Cowan, 2010; Zhang & Luck, 2009;). However, the opportunity to engage in consolidation counteracts this process, rendering the memoranda less vulnerable to ongoing interference (Knöchel et al., 2015; Nieuwenstein & Wyble, 2014). Ricker and Cowan (2010) observed that retention of three unfamiliar visual characters[1], presented in a single array, declined with increasing retention interval. However, in a follow up, Ricker and Cowan (2014) observed that extending the interval between the presentation of each visual stimulus counteracted the impact of an increasing retention interval by enabling more time to engage in consolidation. In contrast to encoding, consolidation of a memory trace continues even after the stimulus is no longer present, occurring until attention is diverted away from the memory trace due to the presentation of another stimulus or activity (Ricker, 2015).

Within recent reports, the assumption has been that extending the opportunity for consolidation of a memory trace strengthens it further, rendering it even more resistant to any potential source of interference. Of course, historically the temporal gradient of

---

[1] The unfamiliar visual characters in Ricker and Cowan (2010) consisted of Greek, Cyrillic and Arabic letters. The authors argued that this stimulus set was distinct from verbal stimuli of English characters as participants were unfamiliar with these characters and their labels.

retroactive interference – the observation that sources of retroactive interference have maximal impact when presented in close proximity to the memorandum (Brown & Lewandowsky, 2010; Dewar et al., 2014) has been synonymous with the conception of a consolidation process (Wixted, 2005). Jolicœur & Dell'Acqua (1998) present influential data consistent with such an idea. Participants were asked to remember a visual display of 1-3 characters (either letters or symbols) to be remembered for later recall. Following each visual display and a brief post-perceptual mask[2] a secondary task was introduced. This was an auditory choice task in which participants had to judge whether an auditory tone was high or low in pitch. Time available for consolidation was manipulated by varying the onset of the auditory task. They predicted that if consolidation was still ongoing when the auditory task was initiated, the division of attention would result in slowed response times in the secondary task - which is what they found. Others report the same outcome (Jolicœur & Dell'Acqua, 1999; Nieuwenstein & Wyble, 2014; Stevanovski & Jolicœur et al., 2007, 2011 who used simple shapes as memoranda).

The 'flip side' of presumed attentionally demanding consolidation on subsequent tasks is whether limiting the consolidation window will lead to impaired recall in the primary task. Indeed, this too has been demonstrated by Nieuwenstein & Wyble (2014). They demonstrated improved memory recall when a longer consolidation time followed the visual display. In contrast to Jolicœur & Dell'Acqua (1998), they used a higher memory load (using either 4 character or complex Kanji characters) as well as a more difficult colour discrimination secondary task. In doing so, they provide compelling evidence that when faced with high levels of interference, extended time for consolidation is necessary for successful recall.

---

[2] The presentation of a post perceptual mask following a visual display is frequently observed in studies of consolidation as this procedure is believed to end perceptual encoding and trigger consolidation (Ricker, 2015)

Therefore, studies of dual task interference have provided converging evidence that the time course of consolidation can reduce the effect of retroactive interference (Nieuwenstein & Wyble, 2014). This enables the accurate retention and recall of both verbal and visual stimuli. However, the majority of visual studies demonstrating an effect of consolidation have used simple stimuli such as unfamiliar characters (Ricker & Cowan, 2010, 2014) or simple shapes (Nieuwenstein & Wyble, 2014; Stevanovski & Jolicœur et al., 2007, 2011). It is unclear how extended consolidation time may influence the retention of more visually complex stimuli that may be more difficult to retain (Eng, Chen, & 2005; Luria, Sessa, Gotter, Jolicœur & Dell'Acqua, 2010). Therefore, further investigation is needed to determine if the observed benefit of an extended post encoding delay in the present experiments was really driven by more time for consolidation of each memory item. It is sensible, however, to exercise caution – there is great uncertainty as to the time-course of the putative trace strengthening that consolidation implies. Does it plateau? In fact, in the laboratory at Leeds there is some unpublished data (McKeown & Zaksaite, 2019) that as the retention interval is extended (perhaps more than 10 seconds or so) a post-encoding strengthening may then yield to a gradually weakening and more vulnerable trace (these authors observed that memory traces were more vulnerable to a distracting activity when this was introduced late rather than early in the retention interval).

There is, however, another contemporary account of forgetting that might explain the present inter-item spacing benefit. It is possible that increasing the inter-item interval reduces the confusion between memory items on a trial by trial basis (Ecker, Tay et al., 2015; Mercer, 2014). In comparison to the short-term consolidation account, temporal distinctiveness accounts state that successful memory retrieval is determined by the ratio between the inter-item intervals between successive to be remembered items and the retention interval between initial encoding and subsequent retrieval

(Ecker, Tay, et al., 2015; Souza & Oberauer, 2014). It has been argued in the literature that this ratio is critical for successful memory recall because when memory traces are maintained longer in short term memory they are compressed in psychological space (Grange & Cross, 2015) and by reducing the inter-item interval, it produces confusion between those items within their spatial and temporal context (Ecker, Brown, et al., 2015; Ecker & Lewandowsky, 2012).

Yet, with an emphasis on interference created by prior memory traces, distinctiveness accounts should at least partially explain the positive influence of increased inter-item interval in the sorts of task used here. For example, when the number of intervening items is equal, increasing the inter-item interval from 2.5s to 9.7s in the task would be expected to make each item more temporally isolated. As such the amount of proactive interference should be reduced. Various distinctiveness accounts have been proposed since Murdock (1962) but most rely on identifying the source of proactive interference local to an item. Thus, Souza and Oberauer (2014) apply a familiar simple ratio of retention interval (current trial between memory display and probe) and the time that has elapsed since the presentation of the memory display on the previous trial. Without pursuing these idea further here (but see Section 8.6 below) unfortunately inter-item interference cannot account for the absence of difference between 12-short and 4_short in the present experiments.

8.7 What is the memory trace?

Recent research on pattern separation and pattern completion using the BPS is extensive, yet we have only just begun to examine not only the validity of this task but the underlying forgetting mechanisms underpinning these processes. Whilst the present work has revealed the critical influence of inter-item interval in this behavioural task, it has also been demonstrated that it is more than of secondary theoretical interest for forgetting itself – instead, temporal variables clearly underlie the efficiency of pattern separation and pattern completion mechanisms themselves (see next Section). Although the present primary focus has been on the window of opportunity for the encoding of information into memory (and hence pattern separation), naturally the continuous recognition task is one where the processes underlying retrieval (and hence pattern completion) are similarly implicated as participants make their old, new, and similar responses on each trial. In other research reports at Leeds (e.g. Mckeown & Mercer, 2012; Mercer & McKeown, 2014) it has been suggested that, in elaborating the mechanisms underlying encoding and retrieval of specifically non-verbal memory it may be necessary to extend the 'short' in short-term memory. Perhaps the participants in the present experiments relied on some form of relatively longer-term storage (beyond several seconds) of the entire set of stimuli accumulating throughout an experimental session? There are reasons to question this. First, the continuous recognition situation does not demand the sorts of capacity-limited recall of sets of stimuli seen in typical working memory recall tasks; and secondly, if performance was capacity-limited one might expect severe penalties when we extend the number of intervening items beyond 3 or 4 (Cowan, 2011; Hardman & Cowan, 2015), which is not what appears to occur.

Nevertheless, in the present task it is probably prudent to remain undecided as to whether the temporal manipulations are uncovering predominantly short-term memory processes or more long-term mechanisms of retrieval. As may be obvious, the use of the term ‘immediate memory’ in this thesis is intended to avoid any appearance of adhering to conventional short- versus long-term memory distinctions such as in the working memory model of Baddeley and colleagues. The time-scale here may be under a second, several seconds, maybe even minutes. The ‘trace’ is not a ‘sensory’ memory in the usual understanding, but nor is it a ‘context’ memory – a distinction attributed to Durlach & Braida (1969) in their discussion of session-long stimulus memories in auditory intensity discriminations. Although these terms, and this distinction between a sort of sensory trace mode of discriminative responding on the one hand, and a more category-like memory on the other comes out of acoustic theory, the distinction is one that is made elsewhere too. McKeown in various publications has argued that the ‘short’ in short-term memory should be extended to perhaps 20 seconds for auditory information, and also identified a more sensory-trace like mode of memory in both auditory and visual short-term memory situations (e.g. McKeown & Mercer, 2012; McKeown, Mercer, Bugajka, Duffy & Barker, 2019). Such notions of extended ‘sensory-like’ memory traces are made within an exciting time in memory research, where fresh recognition is being given to the way, even without focusing consciously, maintenance of the traces of the immediate past may be constructed and form the background against which changes in the ongoing scene are detected. For example, the mismatch negativity (MMN) elicited in the record of event related potentials when changes in experienced stimuli occur and a preparation usually but not exclusively studied for auditory events, is a phenomenon that some claim is a form of comparator of the most recent chain of events entering memory with new events experienced. Indeed, some have suggested that traditional models of memory in terms both of time of endurance, of attention in encoding and maintenance, may need to be radically revised to deal with longer-term sensory codes that the MMN seem to reveal (Winkler & Cowan, 2005). The important point is that, like the MMN, the ability to both detect ‘regularities’ in the ongoing stream of everyday events versus ‘novelty’ may be the single most exciting feature of tasks like BPS. Again, like the mismatch detector, to identify the new, the memory stream must maintain the familiar.

8.8 The question of bias

The three-response task is a memory discrimination task. Most obviously or most simply, it is testing a person’s sensitivity. For example, it tests whether a participant or observer correctly identifies a stimulus item as one presented earlier; the correct ‘old’ response will be labelled a hit, and conversely reporting ‘new’ will be labelled a false alarm. Now, a function of hits and false alarms offered (Macmillan & Creelman, 2005) to attempt to capture this hit/false alarm behaviour is referred to as an index of memory sensitivity. Perfect sensitivity would be one hundred percent hits and zero false alarms. A common measure capturing this is p(c) which may be adjusted according to the proportions of positive or target items and non-target or lure items in a series of tests trials. What p(c) does not capture, however, is how an observer’s sensitivity may change with changes in their criterion of responding or bias. Indeed one of the biggest challenges facing investigators of discriminative abilities is separating those abilities from criterion bias (Kroll, Yonelinas, Dobbins & Frederick, 2002). The approach adopted by signal detection theorists to consider the relation between hits and false alarms in this sort of old/new memory decision test to compile so-called receiver-operating graphs – plotting the relations between hits and false alarms for a range of different biases to respond (from highly likely termed ‘liberal’ to highly unlikely termed ‘conservative’). The accepted index of sensitivity, which is taken to provide a bias-free measure, arising out of the consideration of such functions is dprime (d’), which brings with it certain fundamental assumptions, notably that the distributions of hits and false alarms are normal (equal probability of occurrence) and have equal variance. The calculation of d’ from raw p(c) scores is now straightforward, with on-line resources available. Like all statistics, what is less straightforward is adhering to the underlying assumptions.

The most important underlying assumption is that observer criterion bias is measurable and therefore discriminability or sensitivity can be separated from the criterion bias. The usual calculations in signal detection theory in a two-alternative forced choice task (for example, between a target signal and a noise) are straightforward. The d’ is the difference between z(hits) and z(false alarms), and criterion bias c is -0.5 times the addition of these terms (although Macmillan & Creelman, 2005, Chapter 2 offer some alternative bias measures). However, consider this statement for the case when 2 alternatives are extended to 3 or more: “The assumption of no bias is not made for theoretical reasons, but rather because it simplifies the model and estimation of its parameters (DeCarlo, 2012, p. 196). This author indeed refers to a similar statement by Green and Swets (1988, p. 409): “Our discussion is limited to the two-alternative forced-choice procedure; the analysis of larger numbers of alternatives is complex and, at this date, has not been accomplished”. The present case of continuous recognition from a class of three stimuli is not m_AFC as modelled by DeCarlo, but the challenge is the same – how to partition response bias across three response distributions, where does one place the decision boundary? The models offered by DeCarlo may be beyond the present thesis (and author) but the insights may be explored, at least for the two-case version of BPS reported here in Chapter 6.

In Chapter 6, Section 6.2.2.2 the criterion bias challenge was confronted for the *slightly* more tractable response profiles for ‘new’ and ‘old’ across the underlying three stimulus identities, new, old and similar. The fact that there are only two responses might make the difficulties of calculation a little easier to consider, but it does not remove the implicit or underlying decision distributions. This is so since a perfect participant would presumably identify new, old and similar items perfectly and their ‘decision space’ will be divided in a tri-partite ‘space’ just like the three-response version of BPS. Critically, one cannot assume this perfect participant is perfectly immune from decision bias, so that the statistical problem is precisely the same as the three-response task. Nevertheless, in Chapter 6 it was supposed that, with one key assumption, criterion c might be calculated, if only on a cautious exploratory basis. The assumption made was that a response of ‘new’ to similar items was correct. This leads to three forms of correct response across the three underlying stimulus types: ‘new’ to new items, ‘old’ to old items and ‘new’ to similar items; and two forms of incorrect response across the three underlying stimulus types: ‘old’ to new items and ‘old’ to new items. Two forms of criterion c calculation were made, one in attempting to capture bias for ‘respond old’ and one for ‘respond similar’. The first for tendency to report old as old, and this is intuitively what one might expect, revealed that of the 21 young participants, 8 showed tended to show a slightly liberal bias to report yes. The second for tendency to report similar as new, in contrast and again along with expectation showed no negative c values – the participants perhaps identified the similar items as familiar and were more conservative in reporting the items as new. Arguably, as noted in Chapter 6 what is really needed is that future investigators produce normative data for BPS to derive meaningful receiver-operating curves – only then can response bias be interpretable for differing stimuli and differing populations.

8.9 Suggestions for future research

8.9.1 What is immediate memory for?

In summary, whether through time for memory consolidation of items (Wixted, 2004) or through enhanced temporal distinctiveness (Brown et al., 2007), the inter-item interval in continuous recognition tasks indexing pattern completion and pattern separation is a vital controlling variable. For cognitively unimpaired younger sample of participants, in which age and dementia related factors have been carefully controlled, inter-item interval appeared to be the critical variable governing encoding/retrieval of brief item memory. Yet many questions remain unanswered. Perhaps paramount is the goal of mapping behaviour on specific tasks such as these used here (whether the three-response or two-response variants) to putative circuitry within hippocampus. It has already been suggested within this thesis that the questions ‘what is pattern separation?’ and ‘what is pattern completion?’ cannot be addressed in isolation. Perhaps a better question – an inescapable one – is ‘what are these processes for?’ The American psychologist Daniel Schacter (numerous, but a helpful essay is provided in: Schacter, 2013) offers an increasingly popular adaptive view of memory – memory of the immediate past provides predictive information that enables us to make effective decisions or actions on what is to come. Of course this is a re-statement of the writings of the investigator mentioned in the opening chapter of this thesis – Olga Vinogradova. In a contribution to a conference proceedings (unfortunately the source has been misled) Vinogradova (2001) commented or listed the many and various functions attributed to hippocampus, whether working memory, spatial learning and navigation, behavioural inhibition, memory consolidation and problem solving; and then suggested that some ‘economy’ was called for. For Vinogradova the hippocampus was a novelty detector – a mechanism achieving comparisons of past information with present sensory input. The forms of immediate memory studied in this present series of experiments conform to this type of description – the encoding of the series provides for recognition of what has gone before, of what might have changed slightly in the environment, and what is entirely novel.

8.9.2 What are the most valid stimuli?

There are several directions the present work might take. One is to consider the stimuli. Liu et al. (2016) have, as already noted, offered a number of criticisms of current work in the literature on ‘pattern separation’ (their search term for a systematic review), and one of the most powerful has to do with the types of pictures typically used. The present stimuli address one concern – the easy adoption of a verbal label permitting maintenance of the memory of items through verbal rehearsal – by using a single class of stimuli with, broadly, a single label (“door”). Future work could refine this, however, by using a class of stimuli that more precisely adhere to a single label. For example, doors in the present stimulus set fall into types of door – garage door, church door, domestic home door. One possibility is pictures of watches (but one would have to be careful the participant did have an interest in horology!); another is pictures of leaves (but one would need to be careful the participant did not have an interest in botany!). Another, arguably better way to refine the stimulus set would be to use abstract visual objects. McKeown et al. (2014) have discussed the advantages of abstract modern-art-like pictures in studying visual working memory, and arguably these will not only more perfectly resist verbal encoding but offer an important additional advantage. This is that abstract never-before-seen (and unfamiliar) abstract images will not cue the retrieval of past associations. This is an obvious attraction in a task seeking to uncover the operation of what is here termed immediate memory – the recovery of the very recent past uncontaminated by longer-term associations.

8.9.3 The three parameters of time

Another future direction concerns three parameters in the task. Some discussion was offered in this thesis of the temporal distinctiveness of items in psychological space, or more simply, the temporal compression of items as they recede into the mental past like Crowder's telegraph poles (Crowder, 1976 offered an analogy that mental events within memory became more crowded as they retreated into the past like pictorial perspective). If the presentation of events in time is envisaged as unfolding in discrete equally-spaced steps (like the seconds of a clock), it would be a mistake to think of their encoding into memory as falling into equal spaced steps too. Instead, mental time is believed to be logarithmic (Brown et al., 2007). Therefore the programming of the items in the three-response task is confounding 'real' time and 'mental' time. The latter is 'cramming' items that occurred earlier in the experimental sequence into a bunch, whilst very recent items have good within-items spacings. Future research could certainly address this confound by programming early items with wide inter-item spacing and, increasingly in the series, have late items (or from the participant's point of view, more recent items) having very much more compressed inter-item spacings. In this way the experimenter should be able to achieve equal spacings within memory or mental time – of course different forms of compression would need to be tried; a good starting point would be logarithmic spacings (for example, to base two), following Brown et al. (2007). So the first parameter is memory compression. The other two parameters are those that have formed the focus of the present experiments, memory retention interval and item spacing. So in summary, the three parameters for study are:

- **Memory compression** (a mental variable determined by time since encoding and number of items encoded)
- **Decay interval** (a mental variable determined by trace decay in the absence of active maintenance *and* a physical variable of item recall time in the experiment)
- **Item spacing** (a physical variable determined by the schedule of item presentation during an experiment

8.9.4 The functional role of the trace in ‘noticing the new’

A further future direction does not have to do with the stimuli or their spacings, but seeks to better understand the role of ageing in the dropping off in performance in noticing changes in the environment. Such work could usefully address whether performance is governed by executive memory functions (the sort at play in complex attentionally-demanding tasks) or more ‘simple’ working memory functions (such as when a single simple task is being performed). At first glance the task in the present experiments is a simple one if we focus in on a single stimulus presentation – the response is straightforward. Arguably though, it is the ‘baggage’ of past trials that bring higher load to the attending participant, invoking higher-order or executive functions. So future work might usefully examine interactions across age, proactive interference and varying load or task demands as they develop across trials. This emphasises the major contribution of this thesis – bringing to the foreground the role of contemporary memory theory in understanding the three-response task. For example, one role that must play a dominant role in future thinking is what memory is for – it is, in current thinking, not for keeping account of past events but rather for interpreting the future (like the comparator or change detector discussed in Winkler & Cowan, 2005). Thus, whatever the future directions research into BPS will take, hopefully it will be guided by - or be aware of - those statements about the functional role of the maintained memory trace of the very recent past outlined in this thesis in ‘noticing the new’.

References

Ally, B. A., Hussey, E. P., Ko, P. C., & Molitor, R. J. (2013). Pattern separation and pattern completion in Alzheimer's disease: Evidence of rapid forgetting in amnestic mild cognitive impairment. *Hippocampus, 23*(12), 1246-1258. doi: 10.1002/hipo.22162

Altmann, E. M., & Gray, W. D. (2002). Forgetting to remember: The functional relationship of decay and interference. *Psychological Science, 13*(1), 27-33. doi: 10.1111/1467-9280.00405

Altmann, E. M., & Schunn, C. D. (2012). Decay versus interference: a new look at an old interaction. *Psychological Science, 23*(11), 1435-1437. doi: 10.1177/0956797612446027

Amaral, D. G., & Witter, M. P. (1989). The three-dimensional organization of the hippocampal formation: a review of anatomical data. *Neuroscience, 31*(3), 571-591. doi: 10.1016/0306-4522(89)90424-7

Andersen, P., Morris, R., Amaral, D., Bliss, T., & O'Keefe, J. (2007). *The hippocampus book.* Oxford university press: New York.

Anderson, M. L., James, J. R., & Kirwan, C. B. (2017). An event-related potential investigation of pattern separation and pattern completion processes. *Cognitive Neuroscience, 8*(1), 9-23. doi: 10.1080/17588928.2016.1195804

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science, 18*(7), 622-628. doi: 10.1111/j.1467-9280.2007.01949.x

Awh, E., & Vogel, E. K. (2008). The bouncer in the brain. *Nature Neuroscience*, *11*(1), 5-6. doi: 10.1038/nn0108-5

Azab, M., Stark, S. M., & Stark, C. E. (2014). Contributions of human hippocampal subfields to spatial and temporal pattern separation. *Hippocampus*, *24*(3), 293-302. doi: 10.1002/hipo.22223

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47-89). Academic press. doi: 10.1016/S0079-7421(08)60452-1

Baddeley, A. D., Hitch, G. J., Quinlan, P. T., Bowes, L., & Stone, R. (2016). Doors for memory: A searchable database. *The Quarterly Journal of Experimental Psychology, 69*(11), 2111-2118. doi: 10.1080/17470218.2015.1087582

Bakker, A., Albert, M. S., Krauss, G., Speck, C. L., & Gallagher, M. (2015). Response of the medial temporal lobe network in amnestic mild cognitive impairment to therapeutic intervention assessed by fMRI and memory task performance. *NeuroItem: Clinical, 7*, 688-698. doi: 10.1016/j.nicl.2015.02.009

Bakker, A., Kirwan, C. B., Miller, M., & Stark, C. E. (2008). Pattern separation in the human hippocampal CA3 and dentate gyrus. *Science, 319*(5870), 1640-1642. doi: 10.1126/science.1152882

Bakker, A., Krauss, G. L., Albert, M. S., Speck, C. L., Jones, L. R., Stark, C. E., ... & Gallagher, M. (2012). Reduction of hippocampal hyperactivity improves cognition in amnestic mild cognitive impairment. *Neuron, 74*(3), 467-474. doi: 10.1016/j.neuron.2012.03.023

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General, 133*(1), 83-100. doi: 10.1037/0096-3445.133.1.83

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(3), 570-585. doi: 10.1037/0278-7393.33.3.570

Barrouillet, P., & Camos, V. (2014). On the proper reading of the TBRS model: reply to Oberauer and Lewandowsky (2014). *Frontiers in Psychology, 5* (1331), 1-3. doi: 10.3389/fpsyg.2014.01331

Bayliss, D. M., Bogdanovs, J., & Jarrold, C. (2015). Consolidating working memory: Distinguishing the effects of consolidation, rehearsal and attentional refreshing in a working memory span task. *Journal of Memory and Language, 81*, 34-50. doi: 10.1016/j.jml.2014.12.004

Bennett, I. J., Huffman, D. J., & Stark, C. E. (2014). Limbic tract integrity contributes to pattern separation performance across the lifespan. *Cerebral Cortex*, *25*(9), 2988-2999. doi: 10.1093/cercor/bhu093

Berman, M. G., Jonides, J., & Lewis, R. L. (2009). In search of decay in verbal short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(2), 317-333. doi: 10.1037/a0014873

Bi, J. (2002). Variance of d' for the same-different method. *Behavior Research Methods, Instruments, & Computers, 34*(1), 37-45. doi: 10.3758/BF03195421

Bjork, R. A., & Whitten, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology, 6*(2), 173-189. doi: 10.1016/0010-0285(74)90009-7

Brainerd, C. J., & Reyna, V. F. (1990). Gist is the grist: Fuzzy-trace theory and the new intuitionism. *Developmental Review, 10*(1), 3-47. doi: 10.1016/0273-2297(90)90003-M

Brickman, A. M., Khan, U. A., Provenzano, F. A., Yeung, L. K., Suzuki, W., Schroeter, H., ... & Small, S. A. (2014). Enhancing dentate gyrus function with dietary flavanols improves cognition in older adults. *Nature Neuroscience, 17*(12), 1798-1803. doi: 10.1038/nn.3850

Brown, G. D., & Lewandowsky, S. (2010). Forgetting in memory models. In S. Della Sala (Eds.), *Forgetting* (pp. 49-75). Hove: Psychology Press.

Brown, G. D., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review, 114*(3), 539-576. doi: 10.1037/0033-295X.114.3.539

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: a network model of the phonological loop and its timing. *Psychological Review, 106*(3), 551-581. doi: 10.1037/0033-295X.106.3.551

Carr, V. A., Castel, A. D., & Knowlton, B. J. (2015). Age-related differences in memory after attending to distinctiveness or similarity during learning. *Aging, Neuropsychology, and Cognition, 22*(2), 1-15. doi: 10.1080/13825585.2014.898735

Chen, J., Cook, P. A., & Wagner, A. D. (2015). Prediction strength modulates responses in human area CA1 to sequence violations. *Journal of Neurophysiology, 114*(2), 1227-1238. doi: 10.1152/jn.00149.2015

Cowan, N. (2011). The focus of attention as observed in visual working memory tasks: Making sense of competing claims. *Neuropsychologia, 49*(6), 1401-1406. doi: 10.1016/j.neuropsychologia.2011.01.035

Cowan, N., Saults, J. S., & Nugent, L. D. (1997). The role of absolute and relative amounts of time in forgetting within immediate memory: The case of tone-pitch comparisons. *Psychonomic Bulletin & Review, 4*(3), 393-397. doi:10.3758/BF03210799

Crowder, R. G. (1976). *Principles of learning and memory*. Hillsdale, NJ: Erlbaum.

DeCarlo, L. T. (2012). On a signal detection approach to m-alternative forced choice with bias, with maximum likelihood and Bayesian approaches to estimation. *Journal of Mathematical Psychology, 56*, 196-207. doi: 10.1016/j.jmp.2012.02.004

Déry, N., Goldstein, A., & Becker, S. (2015). A role for adult hippocampal neurogenesis at multiple time scales: A study of recent and remote memory in humans. *Behavioral Neuroscience, 129*(4), 435-449. doi: 10.1037/bne0000073

Déry, N., Pilgrim, M., Gibala, M., Gillen, J., Wojtowicz, J. M., MacQueen, G., & Becker, S. (2013). Adult hippocampal neurogenesis reduces memory interference in humans: opposing effects of aerobic exercise and depression. *Frontiers in Neuroscience, 7*(66), 1-15. doi: 10.3389/fnins.2013.00066

De Schrijver, S., & Barrouillet, P. (2017). Consolidation and restoration of memory traces in working memory. *Psychonomic Bulletin & Review, 25*(5), 1651-1657. doi: 10.3758/s13423-017-1226-7

Deuker, L., Doeller, C. F., Fell, J., & Axmacher, N. (2014). Human neuroimaging studies on the hippocampal CA3 region–integrating evidence for pattern separation and completion. *Frontiers in Cellular Neuroscience, 8*(64), 1-9. doi: 10.3389/fncel.2014.00064

Deuker, L., Olligs, J., Fell, J., Kranz, T. A., Mormann, F., Montag, C., ... & Axmacher, N. (2013). Memory consolidation by replay of stimulus-specific neural activity. *Journal of Neuroscience, 33*(49), 19373-19383. doi: 10.1523/JNEUROSCI.0414-13.2013

Devitt, A. L., & Schacter, D. L. (2016). False memories with age: Neural and cognitive underpinnings. *Neuropsychologia, 91*, 346-359. doi: 10.1016/j.neuropsychologia.2016.08.030

Dewar M, Alber J, Cowan N, & Della Sala, S. (2007). Forgetting due to retroactive iinterference: A fusion of Muller and Pilzecker's (1990) early insights into everyday forgetting and recent research on anterograde amnesia, *Cortex, 43*(5), 616- 634. doi: 0.1016/S0010-9452(08)70492-1

Dewar M, Alber J, Cowan N, & Della Sala, S. (2014) Boosting long-term memory via wakeful rest: intentional rehearsal is not necessary, consolidation is sufficient. *PLoS ONE 9(*10), e109542. doi: 10.1371/journal.pone.0109542

Doxey, C. R., & Kirwan, C. B. (2015). Structural and functional correlates of behavioral pattern separation in the hippocampus and medial temporal lobe. *Hippocampus, 25*(4), 524-533. doi: 10.1002/hipo.22389

Dudai, Y. (2004). The neurobiology of consolidations, or, how stable is the engram?. *Annual Review of Psychology, 55*, 51-86. doi: 10.1146/annurev.psych.55.090902.142050

Duncan, K., Sadanand, A., & Davachi, L. (2012). Memory’s penumbra: episodic memory decisions induce lingering mnemonic biases. *Science, 337*(6093), 485-487. doi: 10.1126/science.1221936

Durlach, N. I., & Braida, L. D. (1969). Intensity perception. I. Preliminary theory of intensity resolution. *The Journal of the Acoustical Society of America,46*(2B), 372- 383. doi: 10.1121/1.1911699

Ecker, U. K., Brown, G. D., & Lewandowsky, S. (2015). Memory without consolidation: Temporal distinctiveness explains retroactive interference. *Cognitive Science, 39*(7), 1570-1593. doi: 10.1111/cogs.12214

Ecker, U. K., & Lewandowsky, S. (2012). Computational constraints in cognitive theories of forgetting. *Frontiers in Psychology, 3*(400), 1-5. doi: 10.3389/fpsyg.2012.00400

Ecker, U. K., Lewandowsky, S., & Oberauer, K. (2014). Removal of information from working memory: A specific updating process. *Journal of Memory and Language*, *74*, 77-90. doi: 10.1016/j.jml.2013.09.003

Ecker, U. K., Tay, J. X., & Brown, G. D. (2015). Effects of prestudy and poststudy rest on memory: Support for temporal interference accounts of forgetting. *Psychonomic Bulletin & Review, 22*(3), 772-778. doi: 10.3758/s13423-014-0737-8

Edmonds, E. C., Glisky, E. L., Bartlett, J. C., & Rapcsak, S. Z. (2012). Cognitive mechanisms of false facial recognition in older adults. *Psychology and Aging, 27*(1), 54-60. doi: 10.1037/a0024582

Eichenbaum, H. (2000). A cortical-hippocampal system for declarative memory. Nature Reviews *Neuroscience, 1*(1), 41-50. doi: 10.1038/35036213

Eng, H. Y., Chen, D., & Jiang, Y. (2005). Visual working memory for simple and complex visual stimuli. *Psychonomic Bulletin & Review, 12*(6), 1127-1133. doi: 10.3758/BF03206454

Ewers, M., Sperling, R. A., Klunk, W. E., Weiner, M. W., & Hampel, H. (2011). Neuroimaging markers for the prediction and early diagnosis of Alzheimer's disease dementia. *Trends in Neurosciences, 34*(8), 430-442. doi: 10.1016/j.tins.2011.05.005

Farrell, S., Oberauer, K., Greaves, M., Pasiecznik, K., Lewandowsky, S., & Jarrold, C. (2016). A test of interference versus decay in working memory: Varying distraction within lists in a complex span task. *Journal of Memory and Language, 90*, 66-87. doi: 10.1016/j.jml.2016.03.010

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191. doi: 10.3758/BF03193146

Gallo, D. A. (2004). Using recall to reduce false recognition: diagnostic and disqualifying monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30(1),* 120-128. doi: 10.1037/0278-7393.30.1.120

Gallo, D. A., Sullivan, A. L., Daffner, K. R., Schacter, D. L., & Budson, A. E. (2004). Associative recognition in Alzheimer's disease: evidence for impaired recall-to-reject. *Neuropsychology, 18*(3), 556-563. doi: 10.1037/0894-4105.18.3.556

Geinisman, Y., de Toledo-Morrell, L., Morrell, F., Persina, I. S., & Rossi, M. (1992). Structural synaptic plasticity associated with the induction of long-term potentiation is preserved in the dentate gyrus of aged rats. *Hippocampus, 2*(4), 445-456. doi: 10.1002/hipo.450020412

Glenberg, A. M., Bradley, M. M., Kraus, T. A., & Renzaglia, G. J. (1983). Studies of the long-term recency effect: Support for a contextually guided retrieval hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*(2), 231-255. doi: 10.1037/0278-7393.9.2.231

Grange, J. A., & Cross, E. (2015). Can time-based decay explain temporal distinctiveness effects in task switching? *The Quarterly Journal of Experimental Psychology, 68*(1), 19-45. doi:10.1080/17470218.2014.934696

Green, D. M., & Swets, J. A. (1988). *Signal detection theory and psychophysics*. Newport Beach.

Guérard, K., Neath, I., Surprenant, A. M., & Tremblay, S. (2010). Distinctiveness in serial memory for spatial information. *Memory & Cognition, 38*(1), 83-91. doi: 10.3758/MC.38.1.83

Gutchess, A. H., & Schacter, D. L. (2012). The neural correlates of gist-based true and false recognition. *Neuroimage, 59*(4), 3418-3426. doi: 10.1016/j.neuroimage.2011.11.078

Hanseeuw, B., Dricot, L., Kavec, M., Grandin, C., Seron, X. & Ivanoiu, A. (2011). Associative encoding deficits in amnestic mild cognitive impairment: a volumetric and functional MRI study. *Neuroitem*, *56*(3), 1743- 1748. doi: 10.1016/j.neuroitem.2011.030.34

Hardman, K. O., & Cowan, N. (2015). Remembering complex objects in visual working memory: Do capacity limits restrict objects or features? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(2), 325. doi: 10.1037/xlm0000031

Hardt, O., Nader, K., & Nadel, L. (2013). Decay happens: the role of active forgetting in memory. *Trends in Cognitive Sciences, 17*(3), 111-120. doi: 10.1016/j.tics.2013.01.001

Hashimoto, R., Abe, N., Ueno, A., Fujii, T., Takahashi, S., & Mori, E. (2012). Changing the criteria for old/new recognition judgments can modulate activity in the anterior hippocampus. *Hippocampus, 22*(2), 141-148. doi: 10.1002/hipo.20878

Hasselmo, M. E., & Schnell, E. (1994). Laminar selectivity of the cholinergic suppression of synaptic transmission in rat hippocampal region CA 1: computational modeling and brain slice physiology. *Journal of Neuroscience, 14*(6), 3898-3914.

Hasselmo, M. E., & Wyble, B. P. (1997). Free recall and recognition in a network model of the hippocampus: simulating effects of scopolamine on human memory function. *Behavioural Brain Research, 89*(1), 1-34. doi: 10.1016/S0166-4328(97)00048-X

Henderson, J. M., & Hollingworth, A. (2003). Eye movements and visual memory: Detecting changes to saccade targets in scenes. *Perception & Psychophysics, 65*(1), 58-71. doi: 10.3758/BF03194783

Holden, H. M., Toner, C., Pirogovsky, E., Kirwan, C. B., & Gilbert, P. E. (2013). Visual object pattern separation varies in older adults. *Learning & Memory, 20*(7), 358-362. doi: 10.1101/lm.030171.112

Huffman, D. J., & Stark, C. E. (2014). Multivariate pattern analysis of the human medial temporal lobe revealed representationally categorical cortex and representationally agnostic hippocampus. *Hippocampus, 24*(11), 1394-1403. doi: 10.1002/hipo.22321

Hunsaker, M. R., & Kesner, R. P. (2013). The operation of pattern separation and pattern completion processes associated with different attributes or domains of memory. *Neuroscience & Biobehavioral Reviews, 37*(1), 36-58. doi: 10.1016/j.neubiorev.2012.09.014

Jolicœur, P., & Dell'Acqua, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology, 36*(2), 138-202. doi: 10.1006/cogp.1998.0684

Jolicœur, P., & Dell'Acqua, R. (1999). Attentional and structural constraints on visual encoding. *Psychological Research, 62*(2), 154-164. doi: 10.1007/s004260050048

Kim, J., & Yassa, M. A. (2013). Assessing recollection and familiarity of similar lures in a behavioral pattern separation task. *Hippocampus, 23*(4), 287-294. doi: 10.1002/hipo.22087

Kirwan, C. B., Hartshorn, A., Stark, S. M., Goodrich-Hunsaker, N. J., Hopkins, R. O., & Stark, C. E. (2012). Pattern separation deficits following damage to the hippocampus. *Neuropsychologia, 50*(10), 2408-2414. doi: 10.1016/j.neuropsychologia. 2012.06.011

Kirwan, C. B., & Stark, C. E. (2007). Overcoming interference: An fMRI investigation of pattern separation in the medial temporal lobe. *Learning & Memory, 14*(9), 625-633. doi: 10.1101/lm.663507

Knöchel, C., Oertel-Knöchel, V., Bittner, R., Stäblein, M., Heselhaus, V., Prvulovic, D., ... & Linden, D. E. (2015). Consolidation time affects performance and neural activity during visual working memory. *Psychiatry Research: Neuroimaging, 231*(1), 33-41. doi: 10.1016/j.pscychresns.2014.10.025

Koen, J. D., & Yonelinas, A. P. (2014). The effects of healthy aging, amnestic mild cognitive impairment, and Alzheimer's disease on recollection and familiarity: a meta-analytic review. *Neuropsychology Review, 24*(3), 332-354. doi: 10.1007/s11065-014-9266-5

Koutstaal, W. (2003). Older adults encode—but do not always use—perceptual details: Intentional versus unintentional effects of detail on memory judgments. *Psychological Science, 14*(2), 189-193. doi: 10.1111/1467-9280.01441

Koutstaal, W., & Schacter, D. L. (1997). Gist-based false recognition of pictures in older and younger adults. *Journal of Memory and Language, 37*(4), 555-583. doi: 10.1006/jmla.1997.2529

Kroll, N. E. A., Yonelinas, A. P., Dobbins, I. G. & Frederick, C. M. (2002). Separating sensitivity from response bias: Implications of comparisons of Yes-No and Forced-Choice for models and measures of recognition memory. *Journal of Experimental Psychology: General, 131(2)*, 241-254. doi: 10.1037//0096-3445.131.2.241

Lacy, J. W., Yassa, M. A., Stark, S. M., Muftuler, L. T., & Stark, C. E. (2011). Distinct pattern separation related transfer functions in human CA3/dentate and CA1 revealed using high-resolution fMRI and variable mnemonic similarity. *Learning & Memory, 18*(1), 15-18. doi: 10.1101/lm.1971111

Leal, S. L., & Yassa, M. A. (2018). Integrating new findings and examining clinical applications of pattern perception. *Nature Neuroscience*, *21(2)*, 163-173. doi:10.1038/s41593-017-0065-1

Lewandowsky, S., Brown, G. D., Wright, T., & Nimmo, L. M. (2006). Timeless memory: Evidence against temporal distinctiveness models of short-term memory for serial order. *Journal of Memory and Language, 54*(1), 20-38. doi: 10.1016/j.jml.2005.08.004

Lewandowsky, S., Ecker, U. K., Farrell, S., & Brown, G. D. (2012). Models of cognition and constraints from neuroscience: A case study involving consolidation. *Australian Journal of Psychology, 64*(1), 37-45. doi: 10.1111/j.1742-9536.2011.00042.x

Lewandowsky, S., Geiger, S. M., Morrell, D. B., & Oberauer, K. (2010). Turning simple span into complex span: Time for decay or interference from distractors?. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(4), 958-978. doi: 10.1037/a0019764

Lewandowsky, S., & Oberauer, K. (2009). No evidence for temporal decay in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(6), 1545-1551. doi: 10.1037/a0017010

Lewandowsky, S., & Oberauer, K. (2015). Rehearsal in serial recall: An unworkable solution to the nonexistent problem of decay. *Psychological Review, 122*(4), 674-699. doi: 10.1037/a0039684

Lewandowsky, S., Oberauer, K., & Brown, G. D. (2009). No temporal decay in verbal short-term memory. *Trends in Cognitive Sciences, 13*(3), 120-126. doi: 10.1016/j.tics.2008.12.003

Liu, K. Y., Gould, R. L., Coulson, M. C., Ward, E. V., & Howard, R. J. (2016). Tests of pattern separation and pattern completion in humans—A systematic review. *Hippocampus. 26*(6), 705-717. doi: 10.1002/hipo.22561

Loiotile, R. E., & Courtney, S. M. (2015). A signal detection theory analysis of behavioral pattern separation paradigms. *Learning & Memory, 22*(8), 364-369. doi: 10.1101/lm.038141.115

Luce, R. D.(1963). A threshold theory for simple detection experiment. *Psychological Review, 70*(1), 61-79. doi: 10.1037/h0039723

Luria, R., Sessa, P., Gotler, A., Jolicœur, P., & Dell'Acqua, R. (2010). Visual short-term memory capacity for simple and complex objects. *Journal of Cognitive Neuroscience, 22*(3), 496-512. doi: 10.1162/jocn.2009.21214

Lupyan, G. (2008). From chair to" chair": a representational shift account of object labelling effects on memory. *Journal of Experimental Psychology: General, 137*(2), 348-369. doi: 10.1037/0096-3445.137.2.348

Macmillan, N. A., & Creelman, C. D. (1991). *Detection Theory: A User's Guide*. Mahwah, NJ : Lawrence Erlbaum.

Macmillan, N. A., & Kaplan, H. L. (1985). Detection theory analysis of group data: estimating sensitivity from average hit and false-alarm rates. *Psychological Bulletin, 98*(1), 185-199. doi: 10.1037/0033-2909.98.1.185

Makovski, T., & Jiang, Y. V. (2008). Proactive interference from items previously stored in visual working memory. *Memory & Cognition, 36*(1), 43-52. doi:10.3758/MC.36.1.43

Marr, D. (1971). Simple memory: a theory for archicortex. *Philosophical transactions of the Royal Society of London. Series B, 262*(841), 23-81. doi: 10.1098/rstb.1971.0078

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (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. *Psychological Review, 102*(3), 419-457. doi: 10.1037/0033-295X.102.3.419

McGaugh, J. L. (2000). Memory--a century of consolidation. *Science, 287*(5451), 248-251. doi: 10.1126/science.287.5451.248

McKeown, D., Holt, J., Delvenne, J. F., Smith, A., & Griffiths, B. (2014). Active versus passive maintenance of visual nonverbal memory. *Psychonomic Bulletin & Review, 21*(4), 1-7. doi: 10.3758/s13423-013-0574-1

McKeown, D., & Mercer, T. (2012). Short term forgetting without interference. *Journal of Experimental Psychology: Learning, Memory and Cognition, 38*(4), 1057-1068. doi: 10.1037/a0027749

McKeown, D., Mercer, T., Bugajka, K., Duffy, P., & Barker, E. (2019). Enduring proactive interference in visual memory. *Memory & Cognition*, under review.

McKeown, D., & Zaksaite, T. (2019). Weakening memory traces become increasingly vulnerable to proactive interference. Manuscript submitted for publication.

McNaughton, B. L., & Morris, R. G. (1987). Hippocampal synaptic enhancement and information storage within a distributed memory system. *Trends in Neurosciences, 10*(10), 408-415. doi: 10.1016/0166-2236(87)90011-7

Mercer, T. (2014). The loss of short-term visual representations over time: Decay or temporal distinctiveness? *Journal of Experimental Psychology: Human Perception and Performance, 40*(6), 2281-2288. doi: 10.1037/a0038141

Mercer, T. (2015). Wakeful rest alleviates interference-based forgetting. *Memory, 23*(2), 127-137. doi: 10.1080/09658211.2013.872279

Mercer, T., & Duffy, P. (2015). The loss of residual visual memories over the passage of time. *The Quarterly Journal of Experimental Psychology, 68*(2), 242-248. doi: 10.1080/17470218.2014.975256

Mercer, T., & McKeown, D. (2014). Decay uncovered in nonverbal short-term memory. *Psychonomic Bulletin & Review, 21*(1), 128-135. doi: 10.3758/s13423-013-0472-6

Migo, E. M., Montaldi, D., & Mayes, A. R. (2013). A visual object stimulus database with standardized similarity information. *Behavior research methods, 45*(2), 344-354. doi: 10.3758/s13428-012-0255-4

Molitor, R. J., Ko, P. C., Hussey, E. P., & Ally, B. A. (2014). Memory-related eye movements challenge behavioral measures of pattern completion and pattern separation. *Hippocampus, 24*(6), 666-672. doi: 10.1002/hipo.22256

Moran, R., & Goshen-Gottstein, Y. (2015). Old processes, new perspectives: Familiarity is correlated with (not independent of) recollection and is more (not equally) variable for targets than for lures. *Cognitive Psychology, 79*, 40-67. doi: 10.1016/j.cogpsych.2015.01.005

Morcom, A. M. (2015). Resisting false recognition: An ERP study of lure discrimination. *Brain Research, 1624*, 336-348. doi: 10.1016/j.brainres.2015.07.049

Motley, S. E., & Kirwan, C. B. (2012). A parametric investigation of pattern separation processes in the medial temporal lobe. *Journal of Neuroscience, 32*(38), 13076-13084. doi: 10.1523/JNEUROSCI.5920-11.2012

Murdock Jr, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, 64(5), 482 488. doi: 10.1037/h0045106

Myers, C. E., & Scharfman, H. E. (2009). A role for hilar cells in pattern separation in the dentate gyrus: a computational approach. *Hippocampus, 19*(4), 321-337. doi: 10.1002/hipo.20516

Myers, C. E., & Scharfman, H. E. (2011). Pattern separation in the dentate gyrus: a role for the CA3 backprojection. *Hippocampus, 21*(11), 1190-1215. doi: 10.1002/hipo.20828

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., ... & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society, 53*(4), 695-699. doi: 10.1111/j.1532-5415.2005.53221.x

Nieuwenstein, M., & Wyble, B. (2014). Beyond a mask and against the bottleneck: Retroactive dual-task interference during working memory consolidation of a masked visual target. *Journal of Experimental Psychology: General, 143(*3), 1409-1427. doi: 10.1037/a0035257

Norman, K. A., & Schacter, D. L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory & Cognition, 25*(6), 838-848. doi: 10.3758/BF03211328

Oberauer, K., Awh, E., & Sutterer, D. W. (2017). The role of long-term memory in a test of visual working memory: Proactive facilitation but no proactive interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(1), 1-22. doi:10.1037/xlm0000302

Oberauer, K., Farrell, S., Jarrold, C., Pasiecznik, K., & Greaves, M. (2012). Interference between maintenance and processing in working memory: The effect of item–distractor similarity in complex span. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(3), 665-685. doi: 10.1037/a0026337

Oberauer, K., & Kliegl, R. (2006). A formal model of capacity limits in working memory. *Journal of Memory and Language, 55*(4), 601-626. doi: 10.1016/j.jml.2006.08.009

Oberauer, K., & Lewandowsky, S. (2008). Forgetting in immediate serial recall: Decay, temporal distinctiveness, or interference?. *Psychological Review, 115*(3), 544-576. doi: 10.1037/0033-295X.115.3.544

Oberauer, K., Lewandowsky, S.,Farrell,S., Jarrold, C., & Greaves, M. (2012). Modelling working memory: An interference model of complex span. *Psychonomic Bulletin & Review, 19*(5), 779-819. doi: 10.3758/s13423-012-0272-4

Oberauer, K., & Lin, H. Y. (2017). An interference model of visual working memory. *Psychological Review, 124*(1), 21-59. doi: 10.1037/rev0000044

O'Reilly, R. C., & McClelland, J. L. (1994). Hippocampal conjunctive encoding, storage, and recall: avoiding a trade-off. *Hippocampus, 4*(6), 661-682. doi: 10.1002/hipo.450040605

Paige, L. E., Cassidy, B. S., Schacter, D. L., & Gutchess, A. H. (2016). Age differences in hippocampal activation during gist-based false recognition. *Neurobiology of Aging, 46,* 76-83. doi: 10.1016/j.neurobiolaging.2016.06.014

Peich, M. C., Husain, M., & Bays, P. M. (2013). Age-related decline of precision and binding in visual working memory. *Psychology and Aging, 28*(3), 729-743. doi: 10.1037/a0033236

Pertzov, Y., Heider, M., Liang, Y., & Husain, M. (2015). Effects of healthy ageing on precision and binding of object location in visual short term memory. *Psychology and Aging, 30*(1), 26-35. doi: 10.1037/a0038396

Pidgeon, L. M., & Morcom, A. M. (2013). Age-related increases in false recognition: the role of perceptual and conceptual similarity. *Frontiers in Aging Neuroscience, 6,* 283-283. doi: 10.3389/fnagi.2014.00283

Pitarque, A., Meléndez, J. C., Sales, A., Mayordomo, T., Satorres, E., Escudero, J., & Algarabel, S. (2016). The effects of healthy aging, amnestic mild cognitive impairment, and Alzheimer's disease on recollection, familiarity and false recognition, estimated by an associative process-dissociation recognition procedure. *Neuropsychologia, 91,* 29-35. doi: 10.1016/j.neuropsychologia.2016.07.010

Pitarque, A., Meléndez, J., Sales, A., Mayordomo, T., Escudero, J., & Algarabel, S. (2016). Differences in false recollection according to the cognitive reserve of healthy older people. *Aging, Neuropsychology, and Cognition, 23*(5), 625-637. doi: 10.1080/13825585.2016.1146221

Pitarque, A., Sales, A., Meléndez, J. C., & Algarabel, S. (2015). Repetition increases false recollection in older people. *Scandinavian Journal of Psychology, 56*(1), 38-44. doi: 10.1111/sjop.12168

Reagh, Z. M., & Yassa, M. A. (2014). Object and spatial mnemonic interference differentially engage lateral and medial entorhinal cortex in humans. *Proceedings of the National Academy of Sciences, 111*(40), E4264-E4273. doi: 10.1073/pnas.1411250111

Remondes, M., & Schuman, E. M. (2004). Role for a cortical input to hippocampal area CA1 in the consolidation of a long-term memory. *Nature, 431*(7009), 699-703. doi: 10.1038/nature02965

Richler, J. J., Palmeri, T. J., & Gauthier, I. (2013). How does using object names influence visual recognition memory? *Journal of Memory and Language, 68*(1), 10-25. doi: 10.1016/j.jml.2012.09.001

Ricker, T. J. (2015). The role of short-term consolidation in memory persistence. AIMS *Neuroscience, 2* (4), 259-279. doi: 10.3934/Neuroscience.2015.4.259

Ricker, T. J., & Cowan, N. (2010). Loss of visual working memory within seconds: The combined use of refreshable and non-refreshable features. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 36*(6), 1355-1368. doi: 10.1037/a0020356

Ricker, T. J., & Cowan, N. (2014). Differences between presentation methods in working memory procedures: A matter of working memory consolidation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(2), 417-428. doi: 10.1037/a0034301

Ricker, T. J., Cowan, N., & Morey, C. C. (2010). Visual working memory is disrupted by covert verbal retrieval. *Psychonomic Bulletin & Review, 17*(4), 516-521. doi: 10.3758/PBR.17.4.516

Ricker, T. J., Spiegel, L. R., & Cowan, N. (2014). Time-based loss in visual short-term memory is from trace decay, not temporal distinctiveness. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(6), 1510-1523. doi: 10.1037/xlm000001

Ricker, T.J., Vergauwe, E., & Cowan, N. (2016). Decay theory of immediate memory: From Brown (1958) to today (2014). *Quarterly Journal of Experimental Psychology, 69,*1969-1995. doi: 10.1080/17470218.2014.914546

Rolls, E. T. (2007). An attractor network in the hippocampus: theory and neurophysiology. *Learning & Memory, 14*(11), 714-731. doi: 10.1101/lm.631207

Rolls, E. T. (2013). The mechanisms for pattern completion and pattern separation in the hippocampus. *Frontiers in Systems Neuroscience, 7*(74), 1-21. doi: 10.3389/fnsys.2013.00074.

Schacter, D. L. (2013) Memory: Sins and virtues. Annals of the New York Academy of *Sciences, 1303*, 56-60. doi: 10.1111/nyas.12168

Schneider, W., Eschman, A., & Zuccolotto, A. (2002) E-Prime User's Guide. Pittsburgh: Psychology Software Tools Inc. Retrieved from http://www.pstnet.com/eprime.cfm

Sense, F., Morey, C. C., Prince, M., Heathcote, A., & Morey, R. D. (2017). Opportunity for verbalization does not improve visual change detection performance: A state-trace analysis. *Behavior Research Methods, 49*(3), 853-862. doi: 10.3758/s13428-016-0741-1

Shelton, D. J., & Kirwan, C. B. (2013). A possible negative influence of depression on the ability to overcome memory interference. *Behavioural Brain Research, 256,* 20-26. doi: 10.1016/j.bbr.2013.08.016

Shipstead, Z., & Engle, R. W. (2013). Interference within the focus of attention: Working memory tasks reflect more than temporary maintenance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(1), 277-289. doi: 10.1037/a0028467

Smith, T. D., Adams, M. M., Gallagher, M., Morrison, J. H., & Rapp, P. R. (2000). Circuit-specific alterations in hippocampal synaptophysin immunoreactivity predict spatial learning impairment in aged rats. *Journal of Neuroscience, 20*(17), 6587-6593.

Souza, A. S. (2016). No Age Deficits in the Ability to Use Attention to Improve Visual Working Memory. *Psychology and Aging, 31*(5), 456-470. doi: 10.1037/pag0000107

Souza, A. S., & Oberauer, K. (2014). Time-based forgetting in visual working memory reflects temporal distinctiveness, not decay. *Psychonomic Bulletin & Review, 22*(1), 156-162. doi: 10.3758/s13423-014-0652-z

Stark, S. M., Stevenson, R., Wu, C., Rutledge, S., & Stark, C. E. (2015). Stability of age-related deficits in the mnemonic similarity task across task variations. *Behavioral Neuroscience, 129*(3), 257-268. doi: 10.1037/bne0000055

Stark, S. M., Yassa, M. A., Lacy, J. W., & Stark, C. E. (2013). A task to assess behavioral pattern separation (BPS) in humans: Data from healthy aging and mild cognitive impairment. *Neuropsychologia, 51*(12), 2442-2449. doi: 10.1016/j.neuropsychologia.2012.12.014

Stevanovski, B., & Jolicœur, P. (2007). Visual short-term memory: Central capacity limitations in short-term consolidation. *Visual Cognition, 15*(5), 532-563. doi: 10.1080/13506280600871917

Stevanovski, B., & Jolicœur, P. (2011). Consolidation of multifeature items in visual working memory: Central capacity requirements for visual consolidation. *Attention, Perception, & Psychophysics, 73*(4), 1108-1119. doi: 10.3758/s13414-011-0099-5.

Stickgold, R., & Walker, M. P. (2005). Memory consolidation and reconsolidation: what is the role of sleep?. *Trends in Neurosciences, 28*(8), 408-415. doi: 10.1016/j.tins.2005.06.004

Toner, C. K., Pirogovsky, E., Kirwan, C. B., & Gilbert, P. E. (2009). Visual object pattern separation deficits in nondemented older adults. *Learning & Memory, 16*(5), 338-342. doi: 10.1101/lm.1315109

Treves, A., & Rolls, E. T. (1992). Computational constraints suggest the need for two distinct input systems to the hippocampal CA3 network. *Hippocampus, 2*(2), 189-199. doi: 10.1002/hipo.450020209

Treves, A., & Rolls, E. T. (1994). Computational analysis of the role of the hippocampus in memory. *Hippocampus, 4*(3), 374-391. doi: 10.1002/hipo.450040319

Tulving, E., & Markowitsch, H. J. (1998). Episodic and declarative memory: role of the hippocampus. *Hippocampus, 8*(3), 198-204. doi: 10.1002/(SICI)1098-1063(1998)8:3<198::AID-HIPO2>3.0.CO;2-G

Unsworth, N., Heitz, R. P., & Parks, N. A. (2008). The importance of temporal distinctiveness for forgetting over the short term. *Psychological Science, 19*(11), 1078-1081. doi: 10.1111/j.1467-9280.2008.02203.x

Vergauwe, E., Dewaele, N., Langerock, N., & Barrouillet, P. (2012). Evidence for a central pool of general resources in working memory. *Journal of Cognitive Psychology, 24*(3), 359-366. doi: 10.1080/20445911.2011.640625

Vergauwe, E., Hartstra, E., Barrouillet, P., & Brass, M. (2015). Domain-general involvement of the posterior frontolateral cortex in time-based resource-sharing in working memory: An fMRI study. *NeuroItem, 115*, 104-116. doi: 10.1016/j.neuroitem.2015.04.059

Vinogradova, O. S. (2001). Hippocampus as comparator: role of the two input and two output systems of the hippocampus in selection and registration of information. *Hippocampus, 11*(5), 578-598. doi: 10.1002/hipo.1073

Walsh, C. M., Wilkins, S., Bettcher, B. M., Butler, C. R., Miller, B. L., & Kramer, J. H. (2014). Memory consolidation in aging and MCI after 1 week. *Neuropsychology, 28*(2), 273-280. doi: 10.1037/neu0000013

Wang, P., Li, J., Li, H., & Zhang, S. (2013). Differences in learning rates for item and associative memories between amnestic mild cognitive impairment and healthy controls. *Behavioral and Brain Functions, 9*(1), 29-40. doi: 10.1186/1744-9081-9-29

Wickens, T.D. (2002). *Elementary signal detection theory*. Oxford University Press, USA.

Wilson, I. A., Gallagher, M., Eichenbaum, H., & Tanila, H. (2006). Neurocognitive aging: prior memories hinder new hippocampal encoding. *Trends in Neurosciences, 29*(12), 662-670. doi: 10.1016/j.tins.2006.10.002

Wilson, I. A., Ikonen, S., Gallagher, M., Eichenbaum, H., & Tanila, H. (2005). Age-associated alterations of hippocampal place cells are subregion specific. *Journal of Neuroscience, 25*(29), 6877-6886. doi: 10.1523/JNEUROSCI.1744-05.2005

Winkler, I., & Cowan, N. (2005). From sensory to long-term memory: Evidence from auditory memory reactivation studies. *Experimental Psychology, 52*(1), 3-20. doi: 10.1027/1618-3169.52.1.3

Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Annual Review of Psychology, 55*, 235-269. doi: 10.1146/annurev.psych.55.090902.141555

Woodman, G. F., & Vogel, E. K. (2008). Selective storage and maintenance of an object's features in visual working memory. *Psychonomic Bulletin & Review, 15*(1), 223-229. doi: 10.3758/PBR.15.1.223

Yassa, M. A., Lacy, J. W., Stark, S. M., Albert, M. S., Gallagher, M., & Stark, C. E. (2011). Pattern separation deficits associated with increased hippocampal CA3 and dentate gyrus activity in nondemented older adults. *Hippocampus, 21*(9), 968-979. doi: 10.1002/hipo.20808

Yassa, M. A., Mattfeld, A. T., Stark, S. M., & Stark, C. E. (2011). Age-related memory deficits linked to circuit-specific disruptions in the hippocampus. *Proceedings of the National Academy of Sciences, 108*(21), 8873-8878. doi: 10.1073/pnas.1101567108

Yassa, M. A., & Stark, C. E. (2008). Multiple signals of recognition memory in the medial temporal lobe. *Hippocampus, 18*(9), 945-954. doi: 10.1002/hipo.20452

Yassa, M. A., & Stark, C. E. (2011). Pattern separation in the hippocampus. *Trends in Neurosciences, 34*(10), 515-525. doi: 10.1016/j.tins.2011.06.006

Yassa, M. A., Stark, S. M., Bakker, A., Albert, M. S., Gallagher, M., & Stark, C. E. (2010). High-resolution structural and functional MRI of hippocampal CA3 and dentate gyrus in patients with amnestic Mild Cognitive Impairment. *Neuroitem, 51*(3), 1242-1252. doi: 10.1016/j.neuroitem.2010.03.040

Yeung, L. K., Ryan, J. D., Cowell, R. A., & Barense, M. D. (2013). Recognition memory impairments caused by false recognition of novel objects. *Journal of Experimental Psychology: General, 142*(4), 1384-1397. doi: 10.1037/a0034021

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

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233-235. doi: 10.1038/nature06860

Zhang, W., & Luck, S. J. (2009). Sudden death and gradual decay in visual working memory. *Psychological Science*, *20*(4), 423-428. doi: 10.1111/j.1467-9280.2009.02322.x

Zokaei, N., Burnett Heyes, S., Gorgoraptis, N., Budhdeo, S., & Husain, M. (2015). Working memory recall precision is a more sensitive index than span. *Journal of Neuropsychology*, *9*(2), 319-329. doi: 10.1111/jnp.12052

Appendices

Appendix A Analysis of MoCA scores for the main outcome measures from Experiment 3

Table A.1.

Summary of the main outcome measures as a function of lag (4_long, 12_short, 4_short) and MoCA Score (High, Low)

		High MoCA		Low MoCA				95% CL		
	Lag	*M*	*SD*	*M*	*SD*	*t* (20)	*p*	Lower	Upper	*d*
Correct rejection rate	4_long	0.65	0.24	0.50	0.18	1.56	0.14	-0.05	0.35	0.72
	12_short	0.58	0.23	0.48	0.27	0.93	0.37	-0.13	0.33	0.40
	4_short	0.58	0.25	0.49	0.26	0.84	0.41	-0.14	0.33	0.37
Hit rate	4_long	0.61	0.22	0.76	0.14	-1.70	0.11	-0.33	0.03	-0.82
	12_short	0.47	0.22	0.64	0.20	-1.75	0.10	-0.36	0.03	-0.78
	4_short	0.48	0.15	0.73	0.14	-3.71	0.00	-0.38	-0.11	-1.66

		High MoCA		Low MoCA				95% CL		
	Lag	*M*	*SD*	*M*	*SD*	*t* (20)	*p*	Lower	Upper	*d*
False alarm rate	4_long	0.16	0.12	0.21	0.12	-0.83	0.42	-0.15	0.07	-0.37
	12_short	0.16	0.13	0.27	0.18	-1.68	0.11	-0.25	0.03	-0.72
	4_short	0.17	0.17	0.23	0.14	-0.74	0.47	-0.20	0.09	-0.34
Recognition accuracy score (RAS)	4_long	1.48	1.08	1.69	0.79	-0.48	0.63	-1.12	0.70	-0.23
	12_short	1.11	0.98	1.17	0.94	-0.14	0.89	-0.95	0.83	-0.06
	4_short	1.07	0.83	1.49	0.79	-1.18	0.25	-1.18	0.33	-0.53
Pattern separation rate	4_long	0.32	0.18	0.37	0.10	-0.70	0.49	-0.19	0.09	-0.35
	12_short	0.31	0.14	0.29	0.16	0.26	0.80	-0.12	0.15	0.11
	4_short	0.30	0.19	0.26	0.18	0.50	0.62	-0.13	0.21	0.22

		High MoCA		Low MoCA				95% CL		
	Lag	*M*	*SD*	*M*	*SD*	*t* (20)	*p*	Lower	Upper	*d*
Similar bias rate	4_long	0.19	0.16	0.30	0.16	-1.54	0.14	-0.26	0.04	-0.68
	12_short	0.27	0.15	0.25	0.18	0.16	0.87	-0.14	0.16	0.07
	4_short	0.24	0.18	0.28	0.23	-0.47	0.65	-0.22	0.14	-0.20
Behavioural pattern separation (BPS)	4_long	0.58	0.44	0.26	0.42	1.64	0.12	-0.09	0.72	0.73
	12_short	0.15	0.32	0.26	0.69	-0.51	0.62	-0.56	0.34	-0.22
	4_short	0.21	0.52	0.00	0.40	0.99	0.33	-0.24	0.66	0.46
Pattern completion rate	4_long	0.47	0.14	0.43	0.17	0.62	0.54	-0.10	0.18	0.30
	12_short	0.41	0.18	0.42	0.15	-0.11	0.92	-0.17	0.15	-0.05
	4_short	0.38	0.13	0.43	0.19	-0.71	0.49	-0.19	0.09	-0.30

		High MoCA		Low MoCA				95% CL		
	Lag	*M*	*SD*	*M*	*SD*	*t* (20)	*p*	Lower	Upper	*d*
Behavioural Pattern Completion (BPC)	4_long	1.17	0.47	0.45	0.48	3.41	0.00	0.28	1.16	1.50
	12_short	1.09	0.53	0.28	0.21	4.08	0.00	0.39	1.22	2.19
	4_short	0.92	0.41	0.38	0.22	3.43	0.00	0.21	0.87	1.70

Appendix B Further analysis of the pitch judgment task (experiment 4)

The overall mean response accuracy (%) and reaction time (ms) for all the tones presented in the pitch judgment task is presented in Table B.1 . For mean response accuracy, a 2 (tone; high, low) x 2 (consolidation window; short, long) repeated measures ANOVA revealed an effect of consolidation window, $F\ (1, 34) = 13.90$, $p = .001$, $\eta_p^2 = .29$, with greater overall accuracy following a short consolidation window in comparison to a long consolidation window, $t(34) = 3.77$, $p = .001$, $d = .79$. There was no effect of tone, $F\ (1, 34) = .58$, $p = .453$ or interaction between tone and consolidation window, $F\ (1, 34) = .001$, $p = .97$.

The pattern of results were replicated for reaction time when submitted to the same analysis. There was an effect of consolidation window, $F\ (1, 34) = 156.59$, $p < .001$, $\eta_p^2 = .82$, with faster reaction times observed for short consolidation window in comparison to a long consolidation window, $t(34) = 12.51$, $p < .001$, $d = 1.42$. Again, there was no effect of tone, $F\ (1, 34) = .66$, $p = .421$ or interaction between consolation window and tone, $F\ (1, 34) = .01$, $p = .91$.

Table B-1

Mean response accuracy and reaction time for all tones presented in the pitch judgement task as a function of the consolidation window (short, long) and the pitch of the tone presented (low, high). Standard deviation is given in parentheses.

Consolidation window	Tone pitch	Response accuracy (%)	Reaction time (ms)
Long	Low	92.03 (9.00)	617.86 (95.19)
	High	91.62 (9.03)	611.83 (94.45)
	Overall	91.76(8.52)	615.05 (89.61)
Short	Low	96.46 (2.84)	500.49 (80.18)
	High	96.01 (3.04)	493.64 (78.19)
	Overall	96.24 (2.75)	496.92 (76.95)

Appendix C Generating similar intervening item groups for Experiment 6

C.1 Overview

For the continuous recognition task conducted in Experiment 6, the key manipulation was the relationship between the four intervening items separating old and similar item pairs. Specifically, the degree of overlapping or similar features shared between the intervening items was varied in order to create three intervening item conditions; novel (four different unrelated items), repeat (one item repeated four times) and similar (four different items which share similar visual features). To create the similar intervening item group, an online similarity rating study was conducted.

C.2 Method

C.2.1 Participants

Seventeen participants (11 female, *Mage* = 36.06 years, *SDage* = 15.25 years) volunteered to complete an online rating study. All participants were native English speakers with self-reported normal or corrected to normal vision and no prior diagnosis of cognitive impairment. Participants were recruited from the University of Leeds community.

C.2.2 Materials

Visual stimuli consisted of 400 coloured items of doors. Stimuli were divided into groups of four items which were matched according to the following criteria, function, age, colour, glazing, condition, shape, opening, details, surround and memorability. Each item groups was arranged in a 2x 2 formation and measure height of 227 pixels and width of 178 pixels. Similarity ratings were collected online using the software, Qualtrics (2016; www.qualtrics.com)

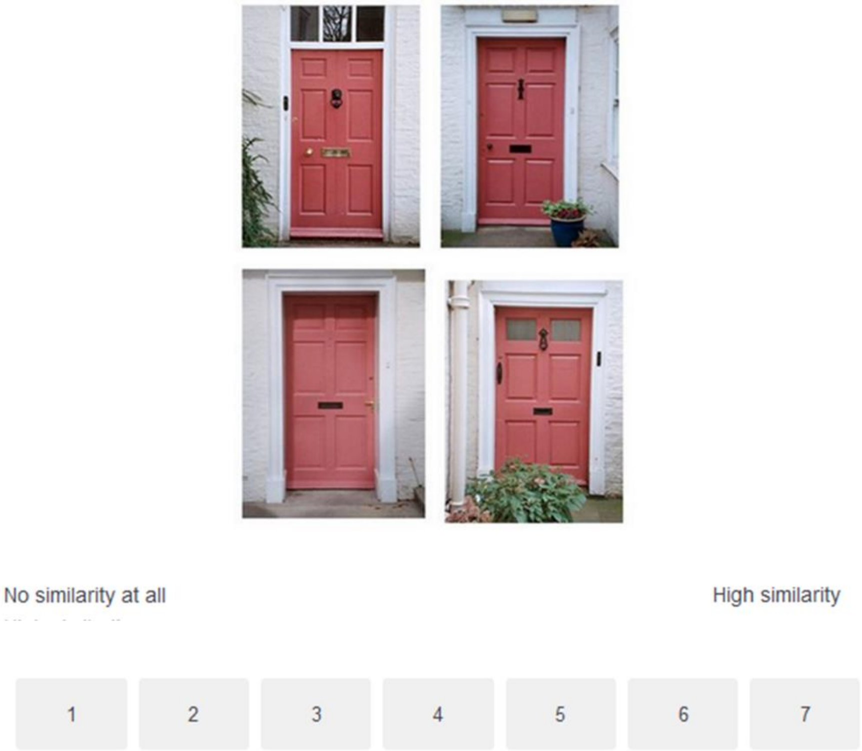

Figure C.1. Example of the similarity rating scale and similar intervening items group

C.2.3 Design

The online similarity task consisted of two parts. Firstly, participants viewed each item group, one at a time. Participants were instructed to take their time to familiarise themselves with each group, as well as the overall degree of similarity across the different groups. Following this, each item group was presented for a second time. For each group, participants were asked to rate the perceptual similarity of the four individual items presented on screen, using a 7 point Likert scale, with 1 representing ‘no similarity at all’ and 7 representing ‘high similarity’. Participants had as long as they needed to provide their responses. The presentation of the item groups was randomised and a prompt to take a 10 minute break was presented after every 10th item group.

Table C.1.

Mean (and standard deviation) similarity for each group of images.

Similarity	*n*	*M*	*SD*	α
1	2	2.74	1.50	0.54
2	11	3.07	1.52	0.91
3	21	3.80	1.49	0.94
4	14	4.19	1.39	0.88
5	22	4.75	1.35	0.89
6	20	5.40	1.29	0.93
7	10	6.35	1.07	0.86

C.3 Results and discussion

Similarity ratings were averaged across participants to calculated the mean and mode similarity rating for each item group. Across the 100 item groups, the mean similarity rating was 4.54 with a standard deviation of 1.67. Table C.1 shows the mean ratings for each similarity group. For the continuous recognition task, 20 similar item groups were required. To identify the similar item groups suitable for the aim of Experiment 6, item groups were ordered and ranked from highest to lowest similarity, based on the mean and mode rating. 20 item groups ranking the highest in terms of similarity were selected from the similar intervening item condition. For the similar intervening items, the mode similarity rating was 7 (*M* = 5.89, *SD* = 1.26). Table C.2 shows the full results of the similarity rating study with the groups selected for the similar intervening item condition highlighted in bold. However, it should be noted that despite ranking highly in terms of similarity, eight similar item groups were not selected as they contained items used as item pair (item group 21, 51, 69, 71, 73, 78, 83 and 100).

Table C.2.

Mode and Mean similarity rating on a scale of 1-7 for each potential similar intervening item group. The similar item groups used in the continuous recognition task are highlighted in bold.

	Mode	M	SD
Group 1	5	5.47	0.62
Group 2	4	3.94	1.48
Group 3	5	4.47	1.66
Group 4	4	4.82	1.42
Group 5	3	4	1.5
Group 6	2	2.94	1.52
Group 7	5	4.71	1.36
Group 8	3	3.59	1.12
Group 9	2	2.88	1.41
Group 10	**6**	**4.65**	**1.66**
Group 11	**6**	**6.35**	**0.61**
Group 12	5	5.12	1.22
Group 13	4	3.71	1.69
Group 14	3	2.53	1.07
Group 15	5	4	1.41
Group 16	2	3.65	1.62
Group 17	4	4.41	1.23
Group 18	2	3.35	1.62

	Mode	M	SD
Group 19	2	3.76	1.39
Group 20	**6**	**6.18**	**0.73**
Group 21	6	5.29	1.1
Group 22	4	4.82	1.19
Group 23	3	3	1.41
Group 24	3	3.29	1.36
Group 25	3	4.12	1.36
Group 26	5	3.82	1.33
Group 27	5	3.76	1.35
Group 28	6	4.65	1.5
Group 29	5	3.88	1.41
Group 30	5	4.82	0.73
Group 31	3	3.88	1.27
Group 32	4	3.88	1.36
Group 33	5	4.71	1.45
Group 34	1	2.53	1.28
Group 35	4	3.59	1.46
Group 36	5	4.29	1.45
Group 37	3	3.53	1.74
Group 38	4	4.24	1.2
Group 39	5	4.65	1.5
Group 40	3	4.12	1.45
Group 41	4	3.29	1.31
Group 42	5	5.24	1.2

	Mode	M	SD
Group 43	2	2.82	1.81
Group 44	4	3.53	1.23
Group 45	2	2.59	1.5
Group 46	3	3.82	1.59
Group 47	5	4.94	0.97
Group 48	5	5	1.17
Group 49	4	4.76	1.03
Group 50	3	3.29	1.4
Group 51	6	5.35	1.11
Group 52	**7**	**6.53**	**1.07**
Group 53	**7**	**6.53**	**0.72**
Group 54	5	6	0.87
Group 55	**7**	**6.47**	**0.72**
Group 56	**7**	**6.94**	**0.24**
Group 57	**7**	**6.59**	**0.71**
Group 58	4	5.12	1.17
Group 59	**6**	**5.94**	**0.9**
Group 60	5	4.53	1.33
Group 61	3	4	1.73
Group 62	3	4.59	1.5
Group 63	**6**	**5.12**	**1.17**
Group 64	**6**	**5.88**	**0.86**
Group 65	6	4.29	1.53
Group 66	**6**	**5.06**	**1.82**

	Mode	M	SD
Group 67	**7**	**6.35**	**1**
Group 68	**7**	**5.82**	**1.51**
Group 69	6	4.94	1.34
Group 70	5	4.47	1.28
Group 71	6	5.47	1.07
Group 72	3	3.53	1.33
Group 73	6	5.59	1.12
Group 74	**7**	**6.12**	**1.32**
Group 75	**6**	**4.76**	**1.52**
Group 76	3	4.06	1.39
Group 77	3	4.76	1.44
Group 78	7	6.59	0.71
Group 79	**6**	**5.53**	**1.23**
Group 80	3	4.53	1.42
Group 81	**6**	**5.65**	**1.17**
Group 82	5	5.47	1.18
Group 83	6	5.59	1.18
Group 84	**7**	**5.59**	**1.5**
Group 85	4	4.41	1.23
Group 86	2	3	1.46
Group 87	3	3.88	1.45
Group 88	2	3.24	1.64
Group 89	5	4.82	1.19
Group 90	1	2.94	1.71

	Mode	M	SD
Group 91	2	3.18	1.33
Group 92	2	2.41	1.23
Group 93	3	3.88	1.54
Group 94	3	3.94	1.3
Group 95	3	3.47	1.59
Group 96	5	4.71	1.45
Group 97	5	5.59	0.94
Group 98	**6**	**5.82**	**0.81**
Group 99	4	4.29	1.26
Group 100	6	5.88	0.99