Kent Academic Repository

Salhi, Louisa (2021) ***Factors that affect incidental encoding during retrieval attempts: Effects of reward, retrieval processes and healthy ageing.*** **Doctor of Philosophy (PhD) thesis, University of Kent,.**

Downloaded from

https://kar.kent.ac.uk/87149/ The University of Kent's Academic Repository KAR

The version of record is available from

https://doi.org/10.22024/UniKent/01.02.87149

This document version

UNSPECIFIED

DOI for this version

Licence for this version

CC BY-NC-ND (Attribution-NonCommercial-NoDerivatives)

Additional information

Versions of research works

Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in *Title of Journal* , Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries

If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies).

Factors that affect incidental encoding during retrieval attempts: Effects of reward, retrieval processes and healthy ageing.

Louisa Salhi

A thesis submitted for the degree of Doctor of Philosophy
at the University of Kent, Canterbury.
77, 403 words

School of Psychology

University of Kent

September 2020

Declaration

I, Louisa Salhi, declare that the work presented in this thesis is my own. The work presented is original and completed under the supervision of Dr Zara Bergström. I have not been awarded a degree by submitting the work included in this thesis for a higher degree at any other institution.

Acknowledgements

I would like to thank Dr Zara Bergström for her inspiration, guidance, support. I am very grateful to you for always having time to talk things through and for always being positive, patient and enthusiastic. Thank you for being my mentor on this journey and for helping steer the way.

Thank you to the Memory Lab, with special thanks to Robin Hellerstedt, Akul Satish, Fadi Ifram, Chloe Brunskill, Arianna Moccia and Matt Plummer, and my fellow PhDs for peer support, engaging debates and talks, as well as the laughs we had. With added thanks to David Vogelsang for his valued feedback and enthusiasm.

I could not have survived my coding journey, which I resisted at the start and now really enjoy, without the psychology technical team, thank you in particular to Frank Gasking, John Allen and Adam Britcher. You not only helped me out of tricky situations but you gave me the tools to help myself out of future ones.

Thank you to all the enthusiastic University of the Third Age members, especially the Canterbury Branch Research Group, Rona Hodges, Jennifer Harrop, Diane Billam, Jacky Moore, Karen Evans, Lynda Bonard. You were all there at the start of this journey and I'm grateful to still be in contact with many of you on submission.

Thank you to all the students who worked on my projects and a big thank you to the participants who gave their time for my research. Your contribution was truly invaluable to me. In particular the amazing older adults that took part in my studies that all had truly inspirational stories to tell and guidance to share, you have all helped shape my outlook to life. Those were some of the most fun testing sessions, you made the (long) EEG set up fly by.

I have personal thanks to give to my loving and supportive family and friends. Luckily, I have a large family to share the burden of support. Thank you all for putting up with my stress and for always being at the end of a call for me. But a special thanks goes to my parents who have always taught me that education is the key to success, something which no-one can take away from you and that you carry with you through your life. Thank you for giving me the confidence to embark on this journey.

Lastly, a big thank you goes to my partner Alex. You have always given the strength, space and time to pursue my goals and you have supported me through challenging times. Thank you for the sacrifices you have made that have enabled me to reach this point. I am excited to finally have finished my journey in formal education for this first time in our relationship and I am looking forward to our next life chapter.

Statement of my contribution to knowledge

My research in this thesis has contributed to knowledge both directly through formal publications and presentations (listed below) as well as more informally through community talks and citizen science training. This thesis provides wider knowledge of episodic memory processes and factors. In particular I provide novel findings related to healthy ageing, commonly assumed to be deteriorated in age, as well as factors such as rewards that are popularly associated with memory boosts. I have provided rigorous scientific evidence that both of these factors do not unilaterally affect memory and that they are complex and interacting factors that also interact with more automatic recognition processes, some of which are stable over age, whereas others are changed, but not necessarily deteriorated. I hope that through providing this body of work, the publications that have occurred and will be a direct outcome of this work, alongside the research communications will provide insights into the wider issues around healthy ageing as well as the complexity of episodic memory.

Publications

Experiment 1, in Chapter 2 contributed towards the publication:

Salhi, L., & Bergström, Z. M. (2020). Intact strategic retrieval processes in older adults: no evidence for age-related deficits in source-constrained retrieval. *Memory*, *28*(3), 348-361.

Experiment 4, in Chapter 5 is in preparation to contribute to a minimum of at least one publication, with the first currently in preparation.

Salhi, L., Hellerstedt, R., Vogelsang, D., & Bergström, Z. M. (*in prep*). Electrical Brain activity associated to false familiarity and novelty related increases in subsequent recognition. An exploration into the effect of healthy ageing on recognition processes.

Conferences/Workshops

Kent Memory Malleability workshop (January 10th -11th 2019)

Data blitz talk: "Electrical brain activity associated with false memory-related increases to subsequent recognition." Salhi, L., Moccia, A., Vogelsang, D., & Bergstrom, Z.

Poster presentation: "Electrical brain activity associated with false memory-related increases to subsequent recognition." Salhi, L., Moccia, A., Vogelsang, D., & Bergstrom, Z.

Kent Psychology PhD Annual symposium. (September 18th 2018)

Presentation talk: "Electrical brain activity associated with false memory-related increases to subsequent recognition." Salhi, L & Bergstrom, Z.

British Association of Cognitive Neuroscience Conference at Glasgow University (September 6th -7th 2018)

Poster presentation; "Electrical brain activity associated with false memory-related increases to subsequent recognition." Salhi, L., Moccia, A., Vogelsang, D., & Bergstrom, Z.

Kent Memory Lab internal annual conference (May 9th 2018)

Presentation talk; "Electrical brain activity associated with false memory-related increases to subsequent recognition. Preliminary analysis" Salhi, L., & Bergstrom, Z.

Experimental Psychology Society research meeting at UCL (January 4th 2018)

Poster Presentation; "Factors that Enhance Memory across Age Groups" Salhi, L., U3A Canterbury district research group & Bergstrom, Z.

Postgraduate Research Festival at University of Kent (May 16th 2017)

Poster Presentation; "Factors that affect learning of novel information during retrieval attempts" Salhi, L., Moccia, A., Vogelsang, D., & Bergstrom, Z.

'Stop me if you think you've heard this one before: Novelty, repetition and the brain' Workshop at University of East Anglia. (May 11-12th 2017)

Poster presentation; "Factors that affect learning of novel information during retrieval attempts" Salhi, L., Moccia, A., Vogelsang, D., & Bergstrom, Z.

Non-academic research communications

Public engagement talk– citizen science event organised and held at university of Kent. (October 28th 2017)

Presentation talk; “Factors that Enhance Memory across Age Groups- results, debrief and feedback session” Salhi, L., & Bergstrom, Z.

South East University of the Third Age research conference at Royal Holloway (July 5th 2017)

Poster Presentation; “Factors that Enhance Memory across Age Groups” Salhi, L., U3A Canterbury district research group & Bergstrom, Z.

Research Showcase at University of Kent in partnership with the University of the Third Age (May 23rd 2017)

Presentation talk; “Factors that Enhance Memory across Age Groups; A citizen science collaboration” Salhi, L., Harrop, J., U3A research group & Bergstrom, Z.

Poster Presentation; “Factors that Enhance Memory across Age Groups” Salhi, L., U3A Canterbury district research group & Bergstrom, Z.

Abstract

A successful strategy to aid recognition is to constrain retrieval search towards a specific context in order to facilitate retrieval of goal-relevant memories. The “memory-for-foils” paradigm is used to investigate this process, termed source-constrained retrieval, by assessing whether incidental encoding of new foils during an old/new recognition test differs depending on the type of processing that was previously used during study of the old items in the test. Ageing is thought to involve impairments to cognitive control functions that support episodic memory as well as a reliance on familiarity-based recognition, suggesting that older adults will be less able to constrain retrieval search than younger adults. This thesis extends on prior source-constrained retrieval literature and explored what potential other factors could modulate the rate of incidental foil encoding. Novel factors of external reward, as well as an internal factor of item-specific retrieval processes and subjective judgments were examined, as well as how these factors are affected by healthy ageing. Contrasting to prior findings, here I show that older adults can spontaneously constrain retrieval to the same extent as younger adults, suggesting that ageing-related episodic memory decline, even in cognitively complex tasks, is not inevitable. However, there was little support for external rewards enhancing incidental encoding beyond altering response biases, despite effects of reward on motivation. Item-elicited retrieval processes however were more strongly related to foil memory, with false familiarity and heightened novelty processing for foils during an initial encounter being the most consistent predictors of increased subsequent recognition after incidental encoding. The neural ERP findings also demonstrated that spontaneous recognition processes elicited by new items have a knock-on effect on their incidental encoding, as well as succinctly showing that episodic memory tasks designated as encoding and retrieval phases are not ‘process pure’.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction & Literature Review

Episodic memory enables us as humans to ‘re-live’ experiences from the past, but also to contemplate future experiences based on what we expect to happen and plan for future events. Episodic memory allows us to learn from unique events, which makes the memory system highly adaptive but also vulnerable to change. The current thesis focused on a relatively neglected aspect of episodic memory, which is the issue of how we encode new information while attempting to recognise if we have seen that information before.

In this chapter, I give an overview of the key processes that are relevant to encoding during recognition, as well as the main research paradigms used to study these processes, and the existing body of evidence on this issue. I then focus my review on potential factors that may influence encoding during recognition, but have not been investigated in prior literature. The main factors that were investigated in this thesis are: *healthy ageing*, a known factor that contributes to a decline in episodic memory as people age, and *external rewards*, another known factor that can modulate episodic memory. I also investigated how these variables may interact with *item-specific retrieval processes* that are elicited by particular pieces of information during recognition attempts.

The aim of this thesis is to provide new knowledge of factors that are relevant to everyday life that can influence encoding during recognition attempts. Of particular interest is the ability to boost such encoding, which may in the future help contribute to interventions that can counteract memory decline. As this thesis spans multiple large research areas, I provide an overview of each topic and how they relate in this chapter.

1.1 Overview of episodic memory processes

It has been theorised for more than a century that human memory is not a unitary system but a complex set of processes which interact and affect one another, with clear differences in capacity, retention and level of detail (Squire, 2004; Tulving, 1993). Critical separations are seen between short-term and long-term memory, and within long-term memory between declarative and non-declarative systems (Pause et al., 2013; Squire, Knowlton, & Musen, 1993). The main focus of this thesis lies within the declarative memory systems also known as explicit memory systems because such memories are experienced with conscious awareness. A major distinction within explicit memory is between memories for events—episodic memory—and memory for knowledge or facts—semantic memory (Pause et al., 2013; Squire et al., 1993; Tulving, 1993). The research in this thesis focused on episodic memory with a particular interest in how episodic encoding occurs incidentally during retrieval attempts, and the factors that modulate such encoding.

1.1.1 Episodic Memory

Episodic Memory is the specialised part of the declarative memory system that encodes and retrieves information about events. These memories, unlike semantic memories, can be very rich and detailed, are linked to the self, and typically include details about what happened during the event, when it happened, and where the event took place (Pause et al., 2013). Importantly, episodic memory occurs with one-shot learning, for example after meeting a new colleague just once, you might still be able to subsequently recognise their face, remember their name, as well as event specific details, for example; what you spoke about and where you were when you met them. Other times, you might bump in to your new colleague but only experience a vague sense of recognition, without being able to remember who they are or where you know them from. Thus, episodic memory can be very contextually rich but can also vary substantially in terms of detail, quality and confidence.

As episodic memories allow humans to remember a large number of details about unique events after just one instance, they are a highly important part of the adaptive memory system (Schacter, 1999). One-shot learning poses advantages since it enables humans to rapidly update their knowledge to include recent important information, but this malleability also makes episodic memories vulnerable to inaccurate updating and false memory generation (Finn, 2017; Schacter, 1999). The episodic memory system does not function in isolation however, since it also interacts with the other memory systems, most notably semantic memory, since sematic information is thought to form important “building blocks” in the construction of episodic event-based memories (Pause et al., 2013). Furthermore, it is known that everyday situations, as well as the experimental paradigms used to study memory, are not “process pure” (Richardson-Klavehn & Bjork, 1988). That is, multiple memory systems and processes can be active in any one situation, influencing each other, and contributing to behavioural and brain activity measures of memory.

1.1.2 Episodic Encoding & Retrieval Processes

Episodic memory encoding is the formation of new event memories and the incorporation of new information from events into the existing knowledge or memory store. Episodic memory retrieval refers to the process by which event information that was previously encoded and was subsequently dormant in the memory store is re-activated, leading to a conscious experience of remembering, and making that information available to influence cognition and behaviour. Episodic retrieval is thought to occur when externally or internally generated information (“retrieval cues”) overlap with information in a stored memory trace, enabling that trace to become reactivated (Nyberg, Habib, McIntosh, & Tulving, 2000). Retrieval processing has been studied using a number of different types of tasks, most classically recall and recognition tests (Baddeley, 1990a). 'Recalling' a memory involves generating additional information via memory retrieval that is not present in the cue, whereas recognition involves determining if you have encountered a cue before, often within a specific context. As the research in this thesis was conducted primarily with recognition tests, the following review focuses on evidence of the memory stages and processes involved in recognition.

1.1.3 Incidental and intentional episodic memory processes

Episodic encoding can be effortful and intentional, involving voluntary attempts at storing information in memory, but it can also be involuntary and occur as an unintentional by-product of experience (Baddeley, 1990b). Incidental encoding can occur very automatically with little directed attention. Automatic processing refers to 'bottom-up' processes that are stimulus-driven and do not require cognitive control or conscious decision making, for example to assign attention to an item. It is highly relevant to how we usually encode episodic information in everyday life, where we rarely intentionally try to memorise information about the situations we encounter, and yet our memories for these situations is often strikingly good. There are times however when we know we need to remember critical information, for example for an upcoming test, leading us to engage in intentional encoding attempts. Intentional encoding is a motivated and effortful process involving directing attention to information that needs to be learnt, which may involve the use of specific strategies to maximise learning. Intentional encoding is perhaps most often used in earlier stages in life where education involving high rates of learning is pervasive, whereas incidental encoding has a bigger role in later life (Wagnon, Wehrmann, Klöppel, & Peter, 2019).

The rate of learning of items such as words by incidental compared to intentional encoding is often relatively similar, with some evidence for more successful word recognition after intentional encoding (Craig et al., 2016), whereas incidental encoding success is more variable based on situational factors. If an incidental encoding task is meaningful, engaging and leads to high levels of attention, then subsequent recognition success is typically high, whereas if the encoding task only requires less meaningful and engaging processing, then subsequent recognition is lower (Craik, 2002). This pattern has been demonstrated extensively in the Levels-of Processing literature (Craik, 2002; Craik & Lockhart, 1972), where words that are deeply processed through directing participants to think about their meaning are incidentally encoded at a higher rate compared to words that are shallowly encoded by directing participants to only consider perceptual characteristics of words, such as their sound or letter shape.

Like encoding, episodic retrieval can occur intentionally as a result of goal-directed retrieval attempts, or unintentionally and automatic, leading to the experience that a memory has popped into awareness. Incidental retrieval is of less relevance to the current body of work, but is a common everyday experience and is associated with memory intrusions, which are features of disorders such as depression, anxiety and PTSD where unwanted memories unintentionally intrude on intentional thought (Krans, Näring, Becker, & Holmes, 2009). More relevant here is intentional retrieval, which involves applying effortful memory search processes to try to retrieve desired information from memory, and engagement of additional control and decision processes to use such memory information in the service of goals. Intentional memory retrieval may involve efforts to constrain retrieval search to a particular context in order to reduce the amount of memories to search (Jacoby, Shimizu, Daniels, & Rhodes, 2005), attempts to reinstate contexts that encoding took place in (Karpicke, Lehman,

& Aue, 2014), or attempts to increase directed attention (Dudukovic, DuBrow, & Wagner, 2009), among other strategic techniques.

1.1.4 Episodic recognition

Episodic recognition refers to our ability to recognise stimuli from events we have experienced. As we encounter objects, words, and faces on a regular basis, we undoubtedly recognise a lot of stimuli all the time. For successful episodic recognition however, the episodic memory system therefore often needs to take into consideration the context in which we are recognising items from, to determine if stimuli appeared in a particular event. Multiple factors determine the success of recognition. To ensure successful recognition, a sufficient memory cue should be available, there should be some consistency between the retrieval and the initial encoding processing state or context, but also enough attention needs be allocated for encoding and retrieval (Karpicke et al., 2014; Rugg & Wilding, 2000). This is not to say that all of these factors need to be present, but if they are not recognition is more difficult and subject to increased errors.

1.1.5 Different neurocognitive processes that contribute to episodic recognition

The key theories of episodic recognition memory aim to determine what cognitive processes are responsible and involved in recognising something as previously encountered as opposed to novel (for reviews of recognition theories see Malmberg, 2008; Mandler, 1980; Yonelinas, 2002; Yonelinas, Aly, Wang, & Koen, 2010). In a typical episodic old/new recognition experiment, participants are first presented with items such as words or pictures in a "study phase", that are to be encoded either intentionally or unintentionally (by having participants complete a task that results in incidental encoding). Next, they are given an old/new recognition test where those items from the study ("old") are intermixed with novel ("new") items that were not encountered in that study phase, and participants are asked to indicate which items they recognise from the study. Classical models of how such recognition memory decisions are made draw on Signal Detection Theory, which posits that in order to recognise something as seen before, you must decide if it elicits a sufficiently strong internal memory signal relative to the background noise elicited by the novel items (Wixted, 2007). That is, it is assumed that all old and new items elicit some degree of memory experience, which arises from a combination of signal and noise. Therefore, the memory "strength" experience varies within each category of old vs. new items so that it approximates a normal distribution across items, with the memory experience distribution for "old" items shifted to be stronger than the memory experience distribution for "new" items overall. In order to discriminate old from new items, participants have to apply a threshold that is internally decided (sometimes referred to as criterion) which provides a relative cut off to determine something as old compared to disregarding it as new. If this threshold is completely separating the old from new items because their distributions are non-overlapping, participants will be able to make consistently accurate old responses to old items (hits) and new responses to new items (correct rejections). However, the more the memory strength distributions for old and new items overlap, the more participants will make errors by misclassifying old items as new (misses) and misclassifying new items as old (false alarms). The particular old items that elicit the weakest memory strength will fall on the "wrong" side of the criterion and be misclassified resulting in misses, whereas the particular new items that elicit the strongest memory strength will fall on the other "wrong" side of the criterion and also be misclassified resulting in false alarms.

Any factors that increase the separation between old vs. new memory strength distributions therefore increase the accuracy on old/new recognition tests, for example very efficient encoding of old items. In contrast, factors that decrease the separation between old and new distributions will increase errors, for example if new items are semantically or perceptually similar to old items, they will elicit similar experiences of memory strength as the old items, and therefore it will be difficult to discriminate between old and new items. In the Deese–Roediger–McDermott (DRM) paradigm (Roediger & McDermott, 1995), lures (new items) that are semantically similar to previously seen old items are presented during a recognition task, leading to extremely high false recognition of those lures. Interestingly, confidence for these incorrect false alarms is also high (Roediger & McDermott, 1995) suggesting these are not just erroneous guesses. Instead, these items are recognised based on their semantic overlap with old items, that gives rise to a strong experience of recognition (Gallo & Roediger, 2002; Roediger & McDermott, 1995) that is very automatic and occurs quickly after items are presented (Heit, Brockdorff, & Lamberts, 2004). The DRM phenomena thus demonstrates the interplay between semantic and episodic recognition. Experimentally new items may also be recognised as old because the memory experience they elicit is modulated by extra-experimental experiences (Kinsbourne & George, 1974; Mandler, Goodman, & Wilkes-Gibbs, 1982), or because of their relative frequency in everyday life. This issue is observed with word recognition in particular, since words that are highly frequent in everyday language are more likely to be incorrectly recognised as “old” in an episodic memory test (Mandler et al., 1982).

Importantly, in the Signal Detection Model, individuals may apply such discrimination judgments with differing levels of caution or response bias, by shifting their internal criterion to be either stricter (requiring a stronger memory experience before deciding that an item is old) or laxer (requiring a weaker memory experience before deciding that an item is old). Participants’ response bias can be determined by the relative proportions of false alarm versus misses they make – more false alarms than misses indicates a lax response bias, whereas the opposite pattern indicates a strict response bias. Shifts in the criterion have been linked to the level of importance of the decision at hand. For example, if there are negative consequences at wrongly classifying items as old, you would be more likely to be cautious and classify items as new if uncertain, adopting a stricter criterion to reduce the risk of incorrectly recognising lures as old. Signal Detection Theory provides a good explanation of how recognition judgments are made if they are based on a 'feeling' or gist, which can be explained by the discrimination and criterion elements of the model. There is however debate to how more qualitive, detailed or rich memories are retrieved in response to recognised cues, which led to the need for dual-process models (Wixted, 2007).

It is now widely acknowledged (but still debated) that episodic recognition likely relies on at least two distinct and dissociable processes of memory retrieval: recollection and familiarity (Diana, Reder, Arndt, & Park, 2006; Mandler, 1980; Wixted, 2007; Yonelinas, 2002; Yonelinas et al., 2010). This idea is referred to as the Dual-Process Signal Detection Model. Familiarity is a fast and relatively automatic retrieval process and results in a 'feeling of knowing' that you have encountered an item before. It is argued that familiarity relies on memory strength (Yonelinas, 2002; Yonelinas et al., 2010), and informs recognition decisions in line with the Signal Detection Theory described above. The experience of this feeling can vary from very strong, to much weaker and is thought to be associated to the strength of the memory trace, but does not contribute to the qualitative aspects of the memory experience, such as contextual details about a particular event (Ingram, Mickes, & Wixted, 2012). A common example of familiarity is when you bump into someone in the street you have met before, you might have a very strong feeling of meeting them before but not be able to retrieve any details about where you previously met them, when or why. Whereas recollection on the other hand, is a slower, more intentional process involving retrieval of contextually rich details about the event where an item was previously encountered (Yonelinas, 2002; Yonelinas et al., 2010). Contextual details about an item or experience could be for example: the time of day, location, the smell at the time and even internal context such as your emotions, or your thoughts at the time of the encounter. These are qualitatively rich, and rather than involving a vague feeling of knowing include memorable elements that can be described. Recollection cannot be explained by Signal Detection Theory alone and is argued to be an absolute concept, so that memory retrieval either contains qualitative details about an event or not.

Even though still debated, the 'Dual-Process Signal Detection Model' is a widely accepted theory and has been studied with multiple recognition paradigms. Both behavioural and neuroimaging studies have provided converging evidence in support of a Dual-Process Model of recognition, and multiple versions of these theories have been proposed (Diana et al., 2006; Yonelinas, 2002; Yonelinas et al., 2010). In order to study how different retrieval processes contribute to recognition, researchers have developed novel paradigms to incorporiate measures of recollection that are distinguishable from measures of familiatity. Although standard old/new recognition tasks are useful for estimating participants ability to discriminate between old and new items and have revealed a great deal about factors that influence recognition, such simple tasks are not able to separate different retrieval processes. Therefore, other tasks such as the Remember-Know paradigm are necessary (Donaldson, 1996; Gardiner & Richardson-Klavehn, 2000). In this paradigm, participants are presented with stimuli that have been learnt (intentionally or incidentally), intermixed with novel items in a recognition test that has been modified to allow participant to explicitly decide if they recollect qualitative details about an item (remember), or if they are relying on a feeling of familiarity (knowing) while making recognition decisions (Gardiner & Richardson-Klavehn, 2000). This paradigm is highly useful in providing a rich amount of detail about a recognition judgment, however, it has received criticisms over the lack of clarity for participants in the distinction between the recollective and familiar ratings, and the subjective nature of self-report that prevents verification of recollection accuracy and what memory experience participants attribute to a recollected memory (Taylor & Henson, 2012).

Other methods of investigating recollective responses are source judgments, here participants provide details about the initial “source” of items, that is, in which context that item was previously encountered (Bröder & Meiser, 2007; Johnson, Hashtroudi, & Lindsey, 1993; Mitchell & Johnson, 2000). Being able to retrieve such contextual details is considered a unique characteristic of recollection, and provides an objective measure of participants' recollection accuracy since memory for encoding contexts can be either correct or incorrect (in contrast to the Remember-Know paradigm). Importantly however, it is not a pure measure of recollection, as the specific source information tested may not be recollected, but other non-tested contextual details may be accessible (Taylor & Henson, 2012). Therefore, common versions of source memory tasks typically combine recognition judgements with additional source identification judgements to categorise the source of items, for example, selecting where an item was seen or the background it was shown on.

Confidence judgments have also been added to old/new recognition tasks to enable a richer understanding of the separation of recognition into familiarity and recollection. Recollected memories are commonly associated with high confidence recognition decisions (Yonelinas et al., 2010). It was initially thought that experiences of familiarity would hold lower confidence, but evidence suggests that familiarity leads to a wide range of recognition confidence judgments (Yonelinas et al., 2010), with even very strong and highly confident familiarity recognition judgments being made (Yonelinas, 2001). Within the Signal Detection Model, confidence is expected to be highest when items elicit either a very weak or very strong memory strength/familiarity experience, because those items are easier to discriminate, whereas items that elicit an intermediate familiarity experience fall very close to the decision criterion, and will therefore be associated with lower confidence. Response times also provide insights into recognition judgments and have been seen to map closely onto the dissociation seen with confidence judgments, with decreased response latencies being associated with increased confidence (Weidemann & Kahana, 2016).

Recently emerging evidence suggests that recognition relies on more than two processes, since there is now evidence for the involvement of novelty detection as a separate process that contributes to recognition judgements. Stimuli can be inherently novel if they have never been encountered before, and in infants this type of novelty is high. However, in adults, stimuli novelty is typically low as we have encountered a vast variety of stimuli such as scenes, words, faces and objects, thus the inherent novelty for most stimuli that we can encounter is low. However, stimuli can also be novel if contextually unexpected, since even though the stimuli are not novel, their presence in a particular situation is (Friedman, Cycowicz, & Gaeta, 2001; Nyberg, 2005; Ranganath & Rainer, 2003). Novelty has been considered for a long time to influence recognition simply through the absence of familiarity. In Signal Detection Theory and the Dual-Process Model, when something is novel there is a very weak or no memory trace, and therefore it elicits a memory experience that falls below the criterion threshold for recognition and is rejected as 'new'. Thus, it has been argued that novelty detection is not a process in itself, but the absence of familiarity processing. However, with the advances in neuroimaging we now know that familiarity and novelty are most likely distinct processes that interact and functionally converge, because highly novel stimuli elicit different brain activity patterns than what would be expected based on gradients in familiarity-related brain processing (Daselaar, Fleck, & Cabeza, 2006; Friedman et al., 2001; Kafkas & Montaldi, 2014; Nyberg, 2005; Ranganath & Rainer, 2003). This perspective thus links to the 'novelty encoding hypothesis' which suggests there is enhanced encoding of novel stimuli (Tulving, Markowitsch, Craik, Habib, & Houle, 1996; Yonelinas et al., 2010) as novel items attract attention (Ranganath & Rainer, 2003).

Unlike previously discussed core retrieval functions such as familiarity and recollection that are thought to be directly related to the reactivation of a stored memory trace, post-retrieval monitoring or 'back-end' processes are also thought to be engaged after a recognition decision has been made. Post-retrieval monitoring is the evaluation and deliberation about retrieved information, which might be in regard to the relevance of the retrieved information for the task, or the source of the information, or if the information is in fact a true or erroneous memory. Monitoring is effortful and linked to cognitive control processes (Finley, Tullis, & Benjamin, 2010). Monitoring might lead to additional retrieval attempts, but importantly the amount of information retrieved is not necessarily linked to the amount of monitoring (Henson, Rugg, Shallice, & Dolan, 2000). Interestingly, post-retrieval monitoring also interacts with confidence. Lower confidence has been seen to correspond to increases in post-retrieval monitoring, which is thought to occur because increased uncertainty in the recognition decision requires more evaluation and reassessment (Finn, 2017; Henson et al., 2000). This view is also consistent with increased response times for low confident responses, both of which suggest increased monitoring of recognition judgments. Behaviourally, post-retrieval monitoring is difficult to assess and quantify, however neuroimaging studies have provided interesting insights, with increases in right dorsolateral PFC activity being associated with increased monitoring (Achim & Lepage, 2005; Henson et al., 2000).

1.2 ERP evidence of episodic memory processes

Neuroimaging techniques provide a complementary addition to behavioural measures that are particularly useful for providing insights into multiple neurocognitive processes that may contribute to one behavioural outcome. Along with studies of clinical populations and animals, the recent steep advances in neuroimaging has provided support for the separation of human memory systems, as well as information into the brain structures involved in memory decisions and how these structures interact when encoding and retrieving a memory. Neuroimaging of episodic memory is a vast and rapidly expanding field, so I will restrict this review to the use of Event-Related Potentials (ERPs), which are derived from electroencephalography (EEG), to study episodic memory, since I used the ERP method in the empirical work in this thesis.

1.2.1 ERPs as a tool to study episodic memory

EEG measures the electrical activity on the scalp emitted from the brain and is a well-established method to examine neurocognitive processes (Luck, 2014). EEG, in contrast with other neuroimaging techniques has excellent temporal resolution, to the millisecond (Rugg, Otten, & Henson, 2002), which is advantageous when examining neural processes in real-time, and even when trying to dissociate two processes occurring at similar time-points as topographic distributions of EEG effects can indicate if one or multiple processes are occurring (but this must be interpreted with caution). It should be noted that EEG recording have poor spatial resolution and cannot be used to detect specific sources of electrical activity within the brain (Friedman & Johnson, 2000) unlike other neuroimaging techniques such as functional magnetic resonance imaging (fMRI) or magnetoencephalography (MEG). Therefore, EEG can indicate different processes occurring by examining the amplitude, topography and latency (Friedman & Johnson, 2000) however the electrical activation is only associated to stimuli presentation and related behavioural responses, therefore it is important not to infer or assume causal relationships between electrical brain activity and behavioural responses (Otten & Rugg, 2005).

Event Related Potentials (ERPs) measure time-locked stimuli evoked changes in EEG, for example a neural response to reading a word or hearing a sound. ERPs are directly relatable to behavioural outputs, therefore the neural data can be linked more closely to behaviour than EEG activity that is not time-locked (Friedman & Johnson, 2000). ERPs are generated by averaging across multiple trials that belong to a particular condition (at least 15-30 trials per condition), which improves the signal to noise ratio which can be an issue in EEG studies (Friedman & Johnson, 2000; Picton et al., 2000; Rugg & Allan, 1995). Therefore, the ERPs are only estimates of the overall activity in that condition (for a methods handbook about interpreting ERPs see Otten & Rugg, 2005 and for guidelines about using ERPs to study cognition see Picton et al., 2000). A benefit of EEG is that it is relatively inexpensive and non-invasive meaning it can be used with a wide range of participants comfortably (Friedman & Johnson, 2000). As ERPs are time-locked to an event, some memory tasks such as self-paced free-recall are not suitable for ERP analysis, as it is not possible to time-lock the neural activity in order to create average ERPs unless the exact timing of the event is known. Recognition tasks are therefore very suitable paradigms for ERP research as ERPs are time-lockable to stimuli presentation, often require button presses which cause much less muscle noise than verbally responding, and also importantly trial numbers can be very high. All of these elements provide advantages for using recognition tasks with ERPs ensuring a design that can attain a good signal-to-noise ratio.

In episodic memory studies, EEG is often recorded at different experimental timepoints to capture neural activity during encoding and retrieval stages. ERPs are then often sorted into conditions based on behaviour. In relation to recognition memory, ERP conditions are commonly split into successful recognition (Hits) and incorrect rejection of old items (Miss) as well as conditions for new items being correctly rejected (Correct Rejection - CR) and incorrect recognition of new items (False Alarm - FA). Further separation can be made if examining successful recollection by for example splitting trials based on successful source judgments or based on R/K responses. Successful recognition would indicate that an item has been successfully encoded during the study phase as well as successfully retrieved at test. Other separations can also be made by splitting trials based on encoding manipulation for example, or by comparing groups.

1.2.2 Key Episodic Memory ERP components

ERP components are patterns of brain activity commonly seen in response to stimuli or a behavioural response. Components have a similar pattern in relation to the topography, time course and polarity and functional characteristics. In this section, I will outline the key ERP components relating to recognition memory.

As discussed earlier two of the critical processes involved in episodic recognition are familiarity and recollection. Familiarity is associated with a relatively early components referred to as the *FN400*, *N400* or the *mi-frontal old/new effect*. This is a negative going components that occurs maximally around 300-500ms after stimuli presentation (Friedman & Johnson, 2000; Mecklinger & Jäger, 2012; Rugg & Curran, 2007). Findings have reliably shown that there is a reduction in early negativity for items that are familiar, whereas those that have not been seen before are more negative. Similar early old/new ERP components have been observed for a range of stimuli from familiar and unfamiliar faces (Curran & Hancock, 2007) to pictures (Curran & Cleary, 2003), to words (Curran, 2000) showing that mid-frontal old/new modulations are a robust episodic recognition effect. Moreover, this component often shows a gradual pattern, with non-familiar items showing the most negative potentials, with a gradual decrease in negativity linked to an increased strength of familiarity (Finnigan, Humphreys, Dennis, & Geffen, 2002). Importantly however, items that go on to elicit recollection or correct source memory versus items that do not are not differentiated in this early time-window, as long as they are recognised (Rugg et al., 1998). This suggests that the early familiarity component is quite dissociable from recollection.

However, the theoretical interpretation of these early old/new ERP effects is not completely clear. There are differences in the FN400 and the N400 in terms of the localisation with the FN400 (frontal negativity) being more highly linked to familiarity of items whereas the N400 (central negativity) has been linked to semantic or conceptual priming (Lucas, Taylor, Henson, & Paller, 2012; Rugg & Curran, 2007; Taylor & Henson, 2012; Woollams, Taylor, Karayanidis, & Henson, 2018) and is widely seen in language processing literature. These are highly interlinked and controversially could be the same brain process linking to conceptual processing of repeated stimuli (Curran & Cleary, 2003; Paller, Voss, & Boehm, 2007).

The second component of interest during old/new recognition is associated with recollection. Recollection-related activity can be seen from around 500-800ms after stimuli onset and is a positive going ERP effect (Mecklinger & Jäger, 2012; Rugg & Allan, 1995; Rugg & Curran, 2007). This effect has been coined the *parietal old/new effect* as this effect is positive going and most maximal over the parietal regions (usually more left-lateralised) of the scalp. There is increased positivity for items that elicit a recollection experience (as measured with a variety of methods) versus items that do not. The later timing of this component is consistent with the findings from other methods that recollection of specific details and contextual information is a slower process compared to familiarity. It has been argued that the parietal old/new effect could be more linked to decision making processes, confidence or memory trace strength rather than directly indexing recollection (Finnigan et al., 2002). These accounts are difficult to separate however, since there is convergence in relation to recollective memories having stronger memory traces and holding more confidence. The late parietal old/new effect is moderated by the strength of recollection, with the largest increases in positivity for full recollection compared to weaker, partial recollection (Vilberg, Moosavi, & Rugg, 2006). This supports the view that the parietal old/new effect is also graded in relation to the strength of the recollective experience (Rugg & Allan, 1995; Wilding, 2000). Due to the reliable and replicable findings of temporal and topographic separation of these components, as well as the functional differences between these components, this line of ERP research has generated substantial support for the Dual-Process Theory of recognition (Rugg & Allan, 1995; Woodruff, Hayama, & Rugg, 2006; Yonelinas, 2002).

Familiarity and recollection are not the only processes involved in recognition decisions however. Another process that occurs during recognition decisions relates to novelty of items. The novelty component is called the *P3 novelty effect* (or the *late positive component, P3a or P300 component*) is related to increased positivity in response to heightened novelty. Arguments have been made to suggest that these components are all similar variants of the same component (Polich, 2007). The P3 novelty effect is shown to be maximal around 500-800ms and is positive going for more novel over less novel items. This component is related in part to an orienting response, whereby it seems to relate to orientation of attentional resources to novel or otherwise important stimuli. With repetition, this orienting response decreases and no longer signals that an item is novel (Friedman et al., 2001; Polich, 2007). The P3 effect has been extensively studied in oddball paradigms. In these paradigms, stimuli are presented or tones are played until habituated, then a different stimuli or tone is shown, this is then very novel and different, evoking the P3 orientating response (Debener, Makeig, Delorme, & Engel, 2005). In recognition tests, new items can sometimes elicit novelty P3 effects that can make their ERPs similar or even more positive than those of old items that elicit parietal old/new effects (Herron, Quayle, & Rugg, 2003).

So far, the *FN400*, *late parietal old/new effect* and the *P3 novelty effect* have been discussed in relation to memory retrieval and novelty detection processes, however post-retrieval processes are also important to consider, since they contribute to the behavioural outcome and final recognition decisions. ERP components that are present later into the epoch are linked to such post-retrieval processes. These processes are more related to decision making and monitoring rather than memory reactivation, and are often highly sensitive to the goal or task requirements. One key post-retrieval monitoring ERP component is referred to as the *right frontal old/new effect* or *Late Frontal Effect (LFE)* (Hayama, Johnson, & Rugg, 2008). This effect is often more sustained and occurs later into the epoch compared to the earlier discussed recognition effects, and is localised more frontally, often right lateralised (Wolk et al., 2009). This ERP component was initially thought to reflect retrieval of contextual information or the amount of recognition decisions required, however there is now evidence that this effect relates to more general monitoring and evaluative processes and can be present for both correct or incorrect recognition as evidence found no association to differences in episodic memory decision making, but rather associations with confidence in these decisions (Cruse & Wilding, 2009; Hayama et al., 2008).

In the current thesis, these ERP markers of recognition processes were used to complement behavioural measures of encoding and retrieval in recognition tasks.

1.3 Encoding during Retrieval Attempts

Retrieval is not only a method to access memories, but also a way to adapt and reinforce important or critical information (Schacter, 1999). Thus, increasing evidence suggests that encoding processes are also engaged during retrieval attempts, leading to learning and memory updating. However, memory researchers often attempt to separate episodic encoding and retrieval processes by designating different phases of experiments as “study” and “test”. This task-based separation can mislead findings, for example if researchers are conducting direct comparisons between “study phases” as encoding and “test phases” as retrieval (Buckner, Wheeler, & Sheridan, 2001; Rugg et al., 2002; Stark & Okado, 2003). In contrast, retrieval and encoding processes can be considered as inter-reliant, with a learning event not only involving encoding processes but also memory retrieval processes. Likewise, memory testing does not only involve retrieval processes, since incidental learning also occurs during retrieval attempts (Danckert, MacLeod, & Fernandes, 2011; Roediger & Karpicke, 2006; Rugg et al., 2002; Stark & Okado, 2003). In neuroimaging techniques like fMRI imaging, the lack of retrieval related activity during “test phases” may be due to incidental encoding activity occurring during the same test phase. Therefore, this comparison of activity between “test” vs. “study” phases may lead to misleading ‘null’ effects whereas encoding and retrieval may be co-occurring in both phases (Stark & Okado, 2003).

It is relevant to examine what factors potentially affect incidental encoding specifically during retrieval attempts. Learning during retrieval has not only been observed through neuroimaging techniques but it has also been widely documented through behavioural paradigms, for example by utilising surprise recognition tests for lures or ‘foils’ that were presented in an earlier memory test (Buckner et al., 2001; Nairne, Pandeirada, VanArsdall, & Blunt, 2015), such as in the source-constrained retrieval paradigm (Jacoby, Shimizu, Daniels, et al., 2005), which investigates how strategic retrieval processing is spontaneously engaged and enhances incidental encoding during recognition attempts.

1.3.1 Source-Constrained-Retrieval & Memory-for-foils

Source-Constrained Retrieval (SCR) refers to the idea that people can strategically enhance recognition by focusing their memory search towards a particular encoding context (or “source”). Such constraints on memory search may be achieved by re-implementing processes previously engaged at encoding during subsequent recognition attempts (Jacoby, Shimizu, Daniels, et al., 2005). For example, if an encoding task involved making judgements about the pleasantness of stimuli, then participants may in a subsequent old/new recognition test, focus on whether the stimuli are pleasant or not as a way to elicit recollection of the encoding context. Evidence that participants spontaneously constrain retrieval in this way comes from the "memory-for-foils" paradigm (Jacoby, Shimizu, Daniels, et al., 2005).

In the most relevant version of this paradigm (Jacoby, Shimizu, Velanova, & Rhodes, 2005) participants first encode two different word lists in two different encoding tasks. The first task encourages participants to focus on the meaning of the words by rating them for pleasantness (considered "deep", semantically-oriented processing in the Levels of Processing (LOP) framework; Craik & Lockhart, 1972). In contrast, the second task involves making judgements on a more perceptual, non-semantic basis, such as detecting if the words contain vowels or not (considered "shallow" processing in the LOP framework). Next, recognition of the two lists is tested in two different tests, one where the old words from the deep/semantic encoding task are intermixed with a set of new "foil" words (the "deep test"), and another test where the old words from the shallow/non-semantic encoding task are intermixed with a different set of new foils (the "shallow test"). Participants are informed of which test contains deep versus shallow old items, and it is thought that this knowledge encourages them to spontaneously focus more on semantic information (e.g. the pleasantness of the words) during the deep test than the shallow test, in order to constrain their memory search towards the appropriate source information.

Unsurprisingly, the encoding manipulation typically leads to higher recognition accuracy on the initial deep test than the shallow test, demonstrating a classic LOP effect (Craik & Lockhart, 1972). The novel aspect of the design involves adding a second subsequent surprise test where all the foils from the prior test phases are intermixed with completely new words, and participants are asked to indicate any words they recognise from the experiment. It is typically found that foils that were initially tested with deeply encoded old items are more likely to be later recognised in the second surprise test than foils that were first tested with shallowly encoded old items. This pattern suggests that participants spontaneously reinstate a more semantic processing mode during the deep recognition test than the shallow recognition test, and that doing so leads to enhanced incidental encoding of foils encountered during the deep test.

The SCR effect is well-documented effect and seen multiple times in varying paradigm manipulations (Alban, 2013; Alban & Kelley, 2012; Bergström, Vogelsang, Benoit, & Simons, 2015; Danckert et al., 2011; Dudukovic et al., 2009; Gray & Gallo, 2015; Halamish, Goldsmith, & Jacoby, 2012; Jacoby, Shimizu, Daniels, et al., 2005; Kantner & Lindsay, 2013; Marsh et al., 2009; Stark & Okado, 2003; Vogelsang, Bonnici, Bergström, Ranganath, & Simons, 2016; Vogelsang, Gruber, Bergström, Ranganath, & Simons, 2018; Zawadzka, Hanczakowski, & Wilding, 2017). Other variations beyond LOP manipulations have shown similar spontaneous reinstatement of encoding strategy, such as improved incidental encoding of self-referential items (Bergström et al., 2015), orienting imagination (Danckert et al., 2011), task difficulty (Alban & Kelley, 2012) and survival relevant stimuli (Nairne et al., 2015). These manipulations show that initial encoding processes are spontaneously reinstated to strategically constrain and improve retrieval, with more elaborate, engaged or meaningful processing of items boosting incidental encoding of novel items (Jacoby, Shimizu, Daniels, et al., 2005). Vogelsang et al. (2016, 2018) showed with fMRI that SCR strategies are associated with the spontaneous reinstatement of neurocognitive processes during memory retrieval attempts, which correlates with later behavioural recognition of foils.

There are limited studies to my knowledge that have implemented a memory-for-foils paradigm with ERPs to examine real-time processing of recognition and the incidental encoding of novel items. The most related study used time-frequency analysis of EEG data to examine neuro-cognitive processes underlying incidental encoding during recognition. Vogelsang et al. (2018) examined oscillations related to encoding during recognition, using the memory-for-foils paradigm similar to that of Jacoby, Shimizu, Daniels, et al. (2005), with deep and shallow processing being compared to examine the effects of spontaneously constraining recognition search towards either semantic or non-semantic information, to examine overlap between activation during the initial encoding phase and the recognition test. They found that EEG oscillations were reinstated during source constrained retrieval, with links between decreased alpha oscillations for semantic processing, seen during both the initial encoding as well as the recognition phase. This evidence thus supported the behavioural and fMRI data by indicating that encoding processes are spontaneously re-instated and used to constrain and aid recognition memory search.

However, reinstatement of encoding processes through source-constrained retrieval is not the only factor that influences encoding during retrieval attempts. Other research has shown that subsequent memory for foils is enhanced if a recognition test was conducted under full compared to divided attention, suggesting that attentive processing is necessary for learning during retrieval attempts (Dudukovic et al., 2009). Neuroimaging (fMRI) evidence also shows evidence that encoding processes are engaged during retrieval attempts even when participants are not explicitly encouraged to use source-constrained retrieval activity, with similar encoding-related fMRI activity present during intentional encoding and during incidental encoding of novel items in a recognition test (Buckner et al., 2001; Stark & Okado, 2003), and such brain activity during the initial recognition test predicted subsequent recognition success in a surprise test of novel items. Thus, there could potentially be many factors that influence encoding of novel items during recognition attempts.

In sum, research using the memory-for-foils paradigm has provided converging evidence that the original encoding techniques or strategies used during study phases can be reinstated during recognition tests and determine the rate of incidental encoding during test phases (Alban, 2013; Alban & Kelley, 2012; Danckert et al., 2011; Jacoby, Shimizu, Daniels, et al., 2005). fMRI studies have been used to more broadly with memory-for-foils paradigms to examine the brain mechanisms involved during foil encoding, to better understand how encoding processes are engaged during retrieval attempts (Bergström et al., 2015; Buckner et al., 2001; Stark & Okado, 2003; Vogelsang et al., 2016), but only one study has investigated encoding of foils with EEG oscillations (Vogelsang et al., 2018). Interestingly, there is limited research that has examined what other factors might modulate incidental encoding during memory tests, and what neural components are involved in encoding during recognition attempts using ERPs. Because similar encoding processes seem to be engaged during recognition tests as during typical study phases (Dudukovic et al., 2009; Stark & Okado, 2003), this suggests that factors that influence encoding more generally in other circumstances would also influence encoding during retrieval attempts, and that these factors might be associated with typical encoding and retrieval-related ERPs. In this thesis, I focused on a subset of relevant factors, specifically healthy ageing, reward and salience, and lastly item specific decision-making processes, as reviewed in the following sections.

1.4 Healthy Ageing and Episodic Memory

Ageing is a major life process that affects episodic memory functioning. On first look it is widely accepted that cognitive processes such as attention, memory and even perception are impaired as we age (Cabeza, Anderson, Locantore, & McIntosh, 2002). We can see clear decreases in cognitive function and memory ability on average with increasing age, with more rapid deterioration from the age of 60 (Nyberg, Lövdén, Riklund, Lindenberger, & Bäckman, 2012; Nyberg & Pudas, 2019). However, within healthy older individuals, there are dramatic differences in the onset, rate and extent to which ageing affects decreases in cognitive function and memory (Nyberg & Pudas, 2019), in particular in relation to lifestyle factors (Chan et al., 2018; Josefsson, Xavier De Luna, Pudas, Nilsson, & Nyberg, 2012; Nyberg & Pudas, 2019) and brain or cognitive 'reserve' (Nyberg et al., 2012).

When examining the literature on ageing there are complex overlapping cognitive processes that need to be considered. Deterioration in one process may have knock on effects onto other processes, which then reduce the performance on certain tasks. Importantly however, the process of interest may not be deteriorated, but just the way of measuring this might disproportionately disadvantage the older adult group. In memory research there can be issues with matching younger adult groups to older adults, with younger groups commonly made up of students. Major issues can arise due to a miss match between groups, with undiagnosed cognitive dysfunction and general cognitive ability in older adults (Cabeza et al., 2002), reduced years in education and the level of physical and social activity in older adults (Chan et al., 2018; Josefsson et al., 2012; Nyberg & Pudas, 2019). Another factor may be experimenter effects when younger adults test older adults, the older participants may perform worse as a result of the expectation that older adults have worse memory (Krendl, Ambady, & Kensinger, 2015). These are just some of the factors to consider when discussing healthy ageing and memory performance.

There is widespread acceptance that there is a deterioration of certain memory processes with the preservation of others in older age. Generally, however there is a reduction in both long-term and working memory, with more preserved function relating to habitual, priming and semantic memory (Buckner, 2004). More specifically, it is well documented that older adults have deficits in episodic memory retrieval, in particular free recall (Danckert &

Craik, 2013) and associative memory (Old & Naveh-Benjamin, 2008), however deterioration has not always been seen in old/new recognition tasks (Craik, 1994; Craik & McDowd, 1987; Danckert & Craik, 2013; Fraundorf, Hourihan, Peters, & Benjamin, 2019). This pattern is argued due to free-recall relying on self-generated memory cueing, increased effortful processing, as well as associative memory also relying on a similarly high level of cognitive control processes, known processes that are less effective in older age (Craik & McDowd, 1987; Spencer & Raz, 1995; Taconnat, Clarys, Vanneste, Bouazzaoui, & Isingrini, 2007).

1.4.1 Incidental vs Intentional memory in ageing

The deterioration in episodic memory seen with ageing disproportionately affects intentional uses of memory over unintentional/ incidental encoding and recognition. Intentional encoding and recognition tasks rely heavily on strategic information processing and have been associated with deteriorations in frontal regions in ageing (Old & Naveh-Benjamin, 2008). Older adults have been seen repeatedly to show deficits in memory tasks including intentional encoding compared to encoding after incidental tasks (Wagnon et al., 2019). Ageing-related changes to incidental uses of memory however has attracted less research attention compared to intentional tasks, despite incidental encoding being an important everyday memory function that is predominantly used by older adults over intentional encoding (Wagnon et al., 2019). A recent review by Wagnon et al. (2019) has highlighted the lack of research in this area, but have drawn on the available literature to suggest that incidental encoding is impaired in ageing when there is a reliance on controlled attentive processing that is cognitively strenuous. Level of processing (LOP) manipulations have often been used to examine incidental encoding rates in ageing, and older adults have consistently shown impaired recognition after deep encoding when compared to younger adults, whereas age differences are less notable after shallow encoding. This pattern is argued due to the requirement in deep processing tasks to strategically process items and engage with the meaning of those items, whereas shallow processing tasks result in automatic encoding that is less demanding, but overall results in worse recognition for both age groups. Shallow processing perhaps masks the issues with older adults' memory, as both groups are closer to chance and alterations in response bias can hide potential ageing-related memory deficits. Nevertheless, the review by Wagnon et al. (2019) points out that there are mixed findings in the literature in encoding in ageing and emphasises the need for continued research in this area. Less variation is seen on the old/new recognition tasks between studies especially when the task does not require a high level of cognitive control (see meta-analysis by Old & Naveh-Benjamin, 2008 and by Fraundorf et al., 2019 for similar findings).

1.4.2 Ageing effects on processes engaged during recognition decisions

In recognition tasks, qualitative differences in memory experience between younger and older adults have been observed. Ageing has been associated with an increased reliance on automatic, familiarity-driven recognition based on item memory strength rather than recollection which is much more cognitively demanding (Koen & Yonelinas, 2016; Yonelinas, 2002). Deterioration in source memory ability also supports the notion that recollective processing is particularly impaired by ageing compared to familiarity processing (Fraundorf et al., 2019). Similarly to other tasks that are particularly influenced by ageing, source memory judgements are highly effortful and require a controlled approach, not only recalling the item memories but contextual details which are often only associated incidentally to the item in question. So ageing appears to be particularly associated with a decline in memory for the context or situation that items were encoded in (Spencer & Raz, 1995).

Post-retrieval monitoring is more likely to occur after weak or uncertain retrieval activation and therefore more likely to occur after familiarity-based retrieval. Neural correlates of post-retrieval monitoring have been found to be invariant across age groups (Horne, Koen, Hauck, & Rugg, 2020; Wang, Johnson, De Chastelaine, Donley, & Rugg, 2016) with right frontal activation linked to increased monitoring and evaluative processes. Others however have found age related changes in monitoring (Wegesin, Friedman, Varughese, & Stern, 2002) but as monitoring is highly linked to task demands such as more demanding source judgments, therefore this variability might be accounted for by task complexity.

The reliance on familiarity processes during recognition leads to another common complaint in older adults, an increase in interference from numerous memory traces leading to increases in erroneous or false memory related responses. Older adults show increased susceptibility to false memories (Dodson & Schacter, 2002) with a striking increase in confidence in these inaccurate memories (Berry, Hastings, West, Lee, & Cavanaugh, 2010; Jacoby & Rhodes, 2006). It has been argued that increased rate of false memories in older age occurs when non-elaborate and shallow processing takes place (Jacoby & Rhodes, 2006), old and new items are therefore less distinct when discriminating them from other old or new items, leading to more incorrect responses. Issues with inaccessible source memories also contributes to the increase in false memories with no context to place memories in, with older adults are also being more likely to confuse an imagined event with a true memory (Dodson & Schacter, 2002). However, Jacoby and Rhodes (2006) suggested that increasing the use of elaborate strategies and increasing the engagement with source identification could reduce false memories in older adults, yet, unlike in younger adults, older adults are less likely to spontaneously engage in these types of strategies (Jacoby, Shimizu, Velanova, et al., 2005). Older adults have been seen to have a lower overall confidence in their memory abilities, linking to deficits in self-efficacy (West, Dark-Freudeman, & Bagwell, 2009). Some have argued that the reduced confidence in memory ability negatively impacts the likelihood of strategy use (Berry et al., 2010). Response times are also seen to be consistently slower in older adults (Berry et al., 2010; Salthouse, 2000).

An understanding of the mechanisms behind this ageing-related change in recognition processing have been aided by neuroimaging, which has provided insights into reductions in hippocampal and frontal cortex brain volume in older adults, areas that recollection relies on, whereas more intact brain areas were related to familiarity (Danckert & Craik, 2013). It has been shown that episodic memory decline is related to decline in multiple other cognitive functions (Buckner, 2004; Henson et al., 2016), reinforcing the overlap between episodic memory and other cognitive processes, in particular executive control (Buckner, 2004; Spencer & Raz, 1995). Interestingly however, not all studies have shown a dramatic ageing-related decrease in recollection (Cabeza et al., 2002; Duarte, Ranganath, Trujillo, & Knight, 2006). However, even when behaviour measures of recollective processes are intact, there are neural changes in older adults. Cabeza et al. (2002) for example, have shown that there is a decrease in neural activity in the pre-frontal cortex (PFC) related to low-performing older adults, whereas high performers who could recollect memories, showed bilateral PFC activity. This contrasted with the younger adults as well as low-performing older adults who show lateralised activity. The authors suggest this change in activity patterns with ageing could indicate compensatory mechanisms or processes enabling these high performers to engage in recollection. This supports the views that ageing affects episodic memory in even those older adults who show no marked behavioural decline.

Theories of episodic memory in ageing try to account for these data patterns. A meta-analysis by Fraundorf et al. (2019) discusses a number of different theories relating to impairments in episodic memory with ageing. Some theories related ageing to impairments in processes, such as recollection, as well as reductions in self-initiated mnemonic strategy use, and issues with increased interference and reduced inhibition control, relating to broader issues with control and attention. There are contrasting theories relating more to global deterioration of cognitive processes and other suggesting that older adults rely on semantic compared to episodic information more. Fraundorf et al. (2019) also discuss issues relating to reduced motivation of older adults partaking in memory studies and the issues relating to stereotypes of older adults having worse memory performance. This highlights the variability in approaches to investigating ageing effects on episodic memory, but across these theories there are similar attributes. All theories suggest ageing is related to deficits in controlled and strategic processing. There appears to be a widespread acceptance that ageing results to decreases in recollection over less effortful familiarity processes, and an acceptance that not one process alone is deteriorated, even if not all theories agree with the global deficit approach.

1.4.3 Ageing and memory strategy use

Young adults with intact control processes often self-initiate strategies to enhance the accuracy of their recognition decisions, for example; by strategically orienting their memory search towards a particular context (Jacoby, Shimizu, Daniels, & Rhodes, 2005; Mecklinger, 2010; Rugg & Wilding, 2000), using elaborate processing (Finley et al., 2010), increasing study-time allocations (Ariel, Price, & Hertzog, 2015; Castel, Murayama, Friedman, Mcgillivray, & Link, 2013; Robison & Unsworth, 2017), or using mnemonic techniques such as association or imagery (Finley et al., 2010). Self-initiated strategy use is a key component of successful intentional encoding (Kirchhoff, 2009). However, older adults appear to be less likely than younger adults to self-initiate strategies at encoding and retrieval, or their strategies are less effective (Morcom, 2016), since older adults that report using self-initiated strategies often use shallow techniques such as repetition or increased concentration (Kirchhoff, Anderson, Barch, Jacoby, & Kirchhoff, 2012). The increased reliance on automatic, familiarity-driven recognition rather than recollection is also consistent with a reduced strategic employment of memory processes in older adults (Koen & Yonelinas, 2016; Yonelinas, 2002). However, it has been found that age is not necessarily the best predictor of reductions in self-initiated retrieval strategy use, but executive function is (Taconnat et al., 2007). There have been differences reported in older adults' beliefs in how memory can be controlled cognitively, with a focus on using and practising their memory retrieval, whereas younger adults held a belief more in line with cognitively controlled processes, that they could internally control and use strategies to alter their memory performance (Hertzog, McGuire, Horhota, & Jopp, 2010). Interestingly, training older adults to use mnemonic strategies have indicated that older adults can effectively use strategies once given effective techniques to use (Gross & Rebok, 2011; Kirchhoff et al., 2012; Paxton, Barch, Storandt, & Braver, 2006).

Ageing is also associated with shifts in response bias. On recognition tasks, younger adults are often quite conservative when guessing, whereas older adults are often more liberal, guessing old more than new. This may be due to the reliance on gist-based memory relating more to a reliance on familiarity (Tun, Wingfield, Rosen, & Blanchard, 1998). It has been suggested that this response bias shift might be a mechanism to try to over-compensate for the belief that older adults have a higher 'forgetting' rate, which can cause older adults to experience "stereotype threat" and want to appear forgetful (Fraundorf et al., 2019). Therefore, older adults may have matched recognition hits but increased false alarm rates, leading to worse recognition discrimination performance overall.

As reviewed earlier, source-constrained retrieval is a very strategic and controlled process that young adults can engage to improve recognition test performance. As a recap, this phenomenon refers to when initial encoding strategies are spontaneously reinstated to constrain and restrict memory search, which has the by-product of modulating incidental encoding new items that are shown as foils in the test. As discussed earlier in this section, older adults often show deficits in incidental encoding during deep and elaborate processing. However, to my knowledge only one prior study has investigated the effects of ageing on source-constrained retrieval. This study reported a deficit in spontaneous use of source-constrained retrieval strategies in older adults (Jacoby, Shimizu, Velanova, et al., 2005). Older adults did not show any difference in subsequent recognition accuracy of foils that had previously been shown in the deep versus shallow recognition tests, whereas a young group of participants showed the standard pattern of enhanced memory for deep over shallow foils. This finding was interpreted as evidence that older adults lack the ability to spontaneously constrain their retrieval to a specific source, leading to a reduced effect of incidental encoding. These findings thus converged with prior evidence showing that older adults are less likely to spontaneously use retrieval strategies in memory tasks (Cohn, Emrich, & Moscovitch, 2008; Kirchhoff, Anderson, Barch, & Jacoby, 2012), potentially due to reduced cognitive control and working memory ability. However, this single published paper had very underpowered designs with small sample sizes, and some of the critical ageing x memory condition interactions were not statistically significant. Therefore, the effects of ageing and spontaneous strategy use on incidental encoding rates merits further research, as addressed in the current thesis.

1.4.4 Ageing and ERPs associated with recognition

ERPs provide unique information about the time-course of specific memory processes and how these change over age, unlike with fMRI for example which can provide more insights about structural change and the involvement of different brain structures (Friedman, Nessler, & Johnson, 2007). ERP research on episodic memory comparing older versus younger adults have contributed to theories on how episodic memory changes in older age, although has often shown conflicting and sometimes surprising results.

Behavioural research and theories have long indicated that familiarity processing is spared in healthy ageing, however ERP studies have shown mixed effects. In line with behavioural findings some studies have shown intact familiarity processes (Friedman, 2013; Nessler, Friedman, Johnson, & Bersick, 2007; Nessler et al., 2008). However, some findings have shown diminished *FN400 effects* (Duarte et al., 2006; Guillaume et al., 2009; Wolk et al., 2009) with others showing even absent familiarity ERPs (Wang, de Chastelaine, Minton, & Rugg, 2012). This raises theoretical questions about what the FN400 effect is related to (Paller et al., 2007) as well as issues around recording and analysing ERP's in older adult populations, as familiarity processes appear behaviourally intact.

Recollection is widely accepted to be behaviourally diminished in older adults (Koen & Yonelinas, 2016; Yonelinas, 2002). Similarly to the FN400 effect however, mixed ERP findings on this issue have been found, with some research showing consistencies with behavioural effects and reduced *late parietal old/new effect*s in healthy older adults (Friedman, 2013; Guillaume et al., 2009; Horne et al., 2020; Wolk et al., 2009). Others however have found ERP markers of recollection to be intact in older adults, but only in relation to performance, with only 'high' performers showing intact recollection related activity (de Chastelaine, Mattson, Wang, Donley, & Rugg, 2016; Duarte et al., 2006; Wang et al., 2016) and others showing a reduced or diminished effect only (Wang et al., 2012). Additionally, others have shown intact behavioural recollection but no parietal ERP effect (Horne et al., 2020; Nessler et al., 2008).

However, even though the FN400 and the late parietal old/new effect have shown mixed neural findings, there are numerous studies suggesting that the magnitude of the *P3* *novelty effect* is intact in healthy older adults. Differences have been seen in relation to novelty habituation after repetition, rather than initial novelty responses to single items (Berti, Vossel, & Gamer, 2017; Daffner et al., 2005; Friedman, Kazmerski, & Cycowicz, 1998; Richardson, Bucks, & Hogan, 2011). Yet again there are mixed effects, with some showing decline in P3 effects with decreased amplitude and increased latency effects (Dinteren, Arns, Jongsma, & Kessels, 2014).

In relation to post-retrieval monitoring (*Late frontal effect - LFE*) there are some studies showing similar findings to younger adults whereas others show a lack, an increase, or a cross-over of ERP differences (Wolk et al., 2009). Some suggestions about compensatory mechanisms and increased evaluative processes have been made (Wolk et al., 2009), but overall authors are speculative about the effect of these differences (Friedman, 2013; Friedman et al., 2007; Nessler et al., 2007) due to the range of processes that could be involved in this later time-window. Thus, it is apparent that other factors may play a role in the ability to detect ERP effects related to memory in older age groups, with important considerations needed surrounding memory performance and paradigm requirements.

As demonstrated by the variability seen in the results of ERPs in older adults of ERP components that are very robust in younger adults, there are issues with interpreting ERP's in older adults. As ERPs are made up of averaged single time-locked potentials they are heavily reliant on consistent timing of activity over trials, since the process of averaging across trials will reduce the size of ERP effects that are not time-locked. Older adults' neural activity may be more jittered in time (Cabeza, Nyberg, & Park, 2009; Picton et al., 2000) which created issues when using average ERPs to analysis group differences. Moreover, another timing issue is how time-windows for comparisons across groups are selected. 'Cognitive-slowing' in older adults has been suggested to delay cognitive processes (Cabeza et al., 2009; Friedman et al., 2007; Morcom et al., 2003) meaning that pre-set time-windows can be misleading as these are typically based on reliably time-locked effects in younger adults who may show less variability in latency of the effects.

Task demands and performance factors are also important to consider, matching group performance has been a sticking point in relation to the interpretability of results. Non-matched performance between age groups makes it less clear how to interpret a diminished or lack of ERP effect, if there was also a diminished or lack of a behavioural effect (Rugg & Morcom, 2009). Therefore, matching groups in some way on their episodic memory performance is needed to minimise behavioural confounds. Some studies have split participants based on their education level (Friedman, 2013) and others based on their actual retrieval task performance, and found more similar memory-related ERPs when performance is matched (de Chastelaine et al., 2016; Duarte et al., 2006; Wang et al., 2016). Relatedly, older adult's data may be noisier due to increased blinking and eye-movement artifacts compared to younger adults (Rugg & Morcom, 2009). Therefore, care is needed to try to enhance the ERP signal-to-noise ratio and ensuring EEG pre-processing and artifact removal are consistent across groups to minimise systematic confounds. Thus, ERPs are a useful tool for examining recognition processing across ages, as long as caveats are considered.

In this thesis, I expanded on the literature reviewed thus far to investigate factors beyond source-constrained retrieval that influence encoding during retrieval attempts. I examined factors that are known to modulate episodic memory in other situations, to determine their effect on incidental encoding during retrieval attempts using the memory-for-foils paradigm. As there is such a lack of research on ageing and incidental encoding during retrieval generally, I also investigated how ageing interacts with those relevant factors. Using ERPs in combination with the memory-for-foils task, I also investigated what ERP components occur during initial recognition that affect incidental encoding and therefore subsequent recognition. One such factor is the effects of reward on memory, as reviewed in the next section.

1.5 Reward & Salience effects on Episodic Memory

The memory system relies to some extent, on salience and the ability to prioritise information to be remembered, given that there is too much information at any given time to be committed to memory. Information that is associated with reward is adaptively important, and hence likely to be prioritised in memory. Broadly speaking, rewards are categorised as objects that we have goals to attain or achieve, this can be through time, energy or effort and can have external or internal outcomes (Arias-Carrion, Stamelou, Murillo-Rodriguez, Menendez-Gonzalez, & Poppel, 2010). If external rewards do indeed increase the importance of stimuli, then it is logical to then suggest that performance, learning and in turn memory should be enhanced. This view is deeply rooted in not only the western education system (Hidi, 2016) but also in people's meta-memory beliefs about the importance of information that is remembered over forgotten memories (Rhodes, Witherby, Castel, & Murayama, 2017).

A wide array of reviews have showcased that external rewards lead to an increase in learning and memory (Arias-Carrion et al., 2010; Hidi, 2016; Miendlarzewska, Bavelier, & Schwartz, 2016). Initially research centred around automatic processes expand on animal models of reinforcement learning (Berridge, Robinson, & Aldridge, 2009; Dayan & Balleine, 2002) where certain behaviours that result in rewarding outcomes are learnt and repeated. Animal models have then been applied to human learning and memory processes (Delgado, 2007). There is a clear evolutionary relevance of automatically prioritising or assigning importance to rewarding or valuable information (Berridge et al., 2009; Hidi, 2016; Shohamy & Adcock, 2010). It is highly relevant for our memory systems to be able to highlight and prioritise information which it is important or valuable. The adaptive nature of episodic memory lends itself both to the reinforcement of rewarding behaviours and the avoidance of punishing or negative outcome behaviours (Shigemune, Tsukiura, Kambara, & Kawashima, 2014). Due to the focus of this review around exploring viable factors to enhance incidental learning the focus will only be on positive reward-based learning and not the adverse punishment and avoidance-based learning (see review here if interested; Seymour, Singer, & Dolan, 2007). Research has highlighted the strength of rewards in automatically capturing cognitive processes (Miendlarzewska et al., 2016) such as attention (Anderson, 2016; Kang & Pashler, 2014), effortful processing and monitoring, as well as memory consolidation (Murayama & Kitagami, 2014), which are all known to lead to enhancements in learning and memory.

There are two main theories of why information associated with external rewards show an increase in memory, these are grouped into automatic modulatory effects and into more strategic control of processing (Hidi, 2016; Miendlarzewska et al., 2016). There is still no clear-cut evidence in favour of rewards benefitting learning and memory with mixed findings related to automatic adaptive memory (Adcock, Thangavel, Whitfield-Gabrieli, Knutson, & Gabrieli, 2006; Arias-Carrion et al., 2010; Murty & Adcock, 2014; Shohamy & Adcock, 2010), strategic prioritisation (Castel et al., 2011; Cohen, Rissman, Hovhannisyan, Castel, & Knowlton, 2017; Cohen, Rissman, Suthana, Castel, & Knowlton, 2014) and also negative consequences of rewards, such as the undermining effect (Deci, 1975; Deci, Koestner, & Richard, 2001; Kuhbandner, Aslan, Emmerdinger, & Murayama, 2016). I will discuss the positive effects of rewards in turn and then examine how ageing affects these two systems to see if rewards could be a viable factor in modulating incidental learning in older age. I will delve into this complex topic in relation to how external rewards in particular could modulate episodic memory with the aim to use external rewards as a factor to boost memory performance.

1.5.1 Automatic vs strategic effects of reward salience on episodic memory

The automatic and strategic viewpoints on how reward interacts and modulated memory are not necessarily at odds but I discus two different mechanisms that might be utilised in different situations to optimise memory. This claim does not mean that these two systems cannot interact and cannot be engaged at the same time, however for clarity these will be presented separately.

1.5.2 Automatic reward effects on episodic memory

The first system relates to the automatic effect of rewards on memory. Rewards are seen to cause an adaptive, automatic response in the brain, with the aim to maximise or repeat rewarding, beneficial behaviours. Information that was more important for survival needed to automatically be remembered more compared to and sometimes at the expense of non-important information (Adcock, Thangavel, Whitfield-Gabrieli, Knutson, & Gabrieli, 2006; Berridge & Robinson, 1998; Miendlarzewska, Bavelier, & Schwartz, 2016). This is demonstrated by the adaptive nature of the declarative memory systems (Shohamy & Adcock, 2010). Memories are not rigid, but are malleable and changeable, and critically used to guide future behaviours and goal-seeking actions (Arias-Carrion et al., 2010; Shohamy & Adcock, 2010).

Interestingly, it has been seen that very early, low level visual processes have been seen to be sensitive to rewards. Rewards are inherently important and therefore typically processed as salient, in that they “stand out” from their context and engage preferential processing, such as by capturing attention automatically. When processing reward related information early visual attention is automatically drawn to this salient information (Apitz & Bunzeck, 2012), with evidence more generally supporting the view that rewarding information and anticipatory motivation leads to an automatic bias in attention allocation (Anderson, 2016; Kang & Pashler, 2014; Miendlarzewska et al., 2016; Robinson et al., 2012). This early processing allocation precedes a latter engagement with effortful processing (Ariel et al., 2015). Increases in pupil dilations to rewarding information supports the view that there is an increase in effortful, and perhaps ‘strategic’ increases in processing leading to enhanced memory for these items (Ariel & Castel, 2014; Bijleveld, Custers, & Aarts, 2009; Miendlarzewska et al., 2016).

Pupillary responses and behavioural outcomes indicate differences in cognitive processing; but does not explain how rewards lead to enhancement of memory. In terms of underlying mechanisms, the neurotransmitter dopamine appears central in anticipation and response to rewards and therefore reward-seeking behaviours (Lisman, Grace, & Duzel, 2011; Shohamy & Adcock, 2010). Multiple strands of evidence have highlighted the role of dopamine in goal-seeking behaviours, learning and episodic memory (Arias-Carrion et al., 2010; Hidi, 2016; Miendlarzewska et al., 2016; Shohamy & Adcock, 2010). Dopamine appears to signal the relevance of an event or stimuli, not only being released due to external rewards generally, but due to novelty (Lisman & Grace, 2005; Otmakhova, Duzel, Deutch, & Lisman, 2013), salience, valence (Hauser, Eldar, Purg, Moutoussis, & Dolan, 2019) and surprise (Butterfield & Metcalfe, 2006; Fazio & Marsh, 2009) as well as internally rewarding factors such as intrinsic motivation (DePasque & Tricomi, 2015) including after moments of insight (Kizilirmak, Thuerich, Folta-Schoofs, Schott, & Richardson-Klavehn, 2016) or during states of curiosity (Gruber, Gelman, & Ranganath, 2014).

More recently, it has been shown that dopamine release directly links to the sensitivity in reward-based gains on incidental encoding. Increases in dopamine release were linked to increased sensitivity and selectivity bias towards high value information, resulting in increased incidental encoding (Hauser et al., 2019). The links between dopamine and enhanced episodic memory have been ascribed to increased protein synthesis, synaptic plasticity and hippocampal binding through increases in long-term potentiation (for a review see Lisman et al., 2011). This highlights how integrated the dopaminergic reward system is with the memory system (Hauser et al., 2019; Shohamy & Adcock, 2010; Wittmann et al., 2005). From recent enhancements in neuroimaging we now know that the dopaminergic system is highly functionally connected to the episodic memory systems (Arias-Carrion et al., 2010). Neural models highlight the functional and structural links both through the projection of neurotransmitters but also the physiological connections between the dopaminergic reward circuit and the medial temporal lobes, central to memory formation and storage (Miendlarzewska et al., 2016; Shohamy & Adcock, 2010). In particular, multiple studies have found that rewards during incidental encoding activate both reward circuits and episodic memory systems, which predicts subsequent recognition, and in particular increased recollective experiences during recognition rather than familiarity-based responses (Wittmann et al, 2005; Adcock et al., 2006) and increased binding between items and contexts after intentional encoding (Wolosin, Zeithamova, & Preston, 2012). However, other studies have found brain activity evidence for functional connectivity between reward and episodic memory systems but without behavioural memory enhancement (Wimmer, Braun, Daw and Shahamy, 2014), showing that rewards can induce changes in neurocognitive processing that do not translate into better memory performance.

There have now been a range of studies employing rewards during incidental or intentional encoding, but there is very little evidence on the issue of whether rewards can influence retrieval processing. Elward, Vilberg and Rugg (2015) assessed if reward modulated recollection at the time of testing, by investigating if rewards altered neural activation correlated to accurate source memories. Here they gave participants pictures of objects paired with meaningless coins to study. After the study phases, coin value was revealed and participants had to determine what coin was linked to what picture. Results showed activation of the core recollective network for correctly recognised source items, as well as ventral striatum activation for higher value items. This study indicated that the dopaminergic reward circuit and the memory systems are separate yet show patterns of similar neural activity, for example, with some structures like the mPFC being responsive to both recollection but also to rewards.

As mentioned, the reward circuit does not only respond to physical rewards, but also states of curiosity, novelty and emotional valence. Anticipation of novelty interacts with the dopaminergic reward circuit leading to increases in midbrain activity and episodic memory for novelty cued stimuli, similar to anticipatory reward-based responses (Wittmann, Dolan, & Duzel, 2011). The anticipation and uncertainty of rewards itself is salient and increases dopaminergic firing (Arias-Carrion et al., 2010; Mason, Farrell, Howard-Jones, & Ludwig, 2017) as well as increasing attentional resources for those items (Anderson, 2016). When a reward is expected, it is theorised that reward predictions are made where the likelihood and risk in receiving the reward are weighed up. When there is a discrepancy between the likelihood of rewards and the actual outcome there is a reward prediction error. Higher risk situations means the likelihood of rewards are lower and these situations gives rise to increased encoding (Rouhani, Norman, & Niv, 2018). The argument behind this is that higher risks which result in a rewarding payoff are more salient, surprising and therefore capture attention more than if the reward was available without the level of risk (Jang, Nassar, Dillon, & Frank, 2019).

Related to the above point, providing feedback after recognition judgments has been seen to alter the rate of learning, presumably via mechanisms related to salience, attention and/or prediction errors. It is thought that during retrieval, predictions are made about the accuracy of a recognition decision (Henson & Gagnepain, 2010). When feedback is provided, it is either in line with the brain’s prediction or not, and if the feedback is inconsistent with the prediction then learning is enhanced. Under certain conditions, feedback is more salient and therefore captures attention more than during other conditions. Rewards are one factor that increases the salience of feedback, yet confidence is also another factor. The ‘hypercorrection effect’ is seen reliably within the semantic memory literature and relates to the updating of memory after surprising feedback (Butterfield & Metcalfe, 2006). High confidence errors are seen to be updated and corrected more often than low confidence errors after feedback because the feedback information is surprising, causing a larger predication error to be generated (Butterfield & Metcalfe, 2006; Fazio & Marsh, 2009, 2010).

It is worth noting that there is evidence for the effect of rewards do not increase in a linear manner as reward value increases. Wittmann, Dolan and Duzel (2011) found that there was a non-linear relationship between reward value and memory ability. During a word judgment task non-rewarding or differing levels of reward related feedback was provided. Overall reward-related feedback increased memory, but increases from 20 to 40 cents did not increase the amount of recollected memory. In some cases, the magnitude of rewards however has been seen to alter subsequent recognition. The anticipation of high rewards, compared to low rewards was seen to increase pre-stimulus theta activity that was associated to subsequent increased correct memory and high confidence (Gruber, Watrous, Ekstrom, Ranganath, & Otten, 2013). Suggesting that the anticipation of rewards appears to be just as important as the attainment and feedback related to rewarding outcomes.

1.5.3 Strategic prioritisation of reward-relevant information in episodic memory

Above I have highlighted the importance of an adaptive and malleable memory system, which interacts with the dopaminergic circuits to drive future behaviour and goal-seeking actions as well as to enhance memory for important or valuable information. These characteristics suggest reward processing is very automatic, yet new evidence also suggests an effect on more strategic metacognitive processing and decision-making criteria. The automatic and more controlled processes that are engaged in anticipation or response to rewards are highly interlinked and difficult to separate. I will now review evidence on more controlled and strategic prioritisation of rewards.

The value-directed forgetting literature (Rhodes et al., 2017) investigates people's ability to control what information is retained or forgotten in a strategic manner, based on which information is more vs. less important or valuable. Many studies in this area have investigated if one can control and strategically monitor one's memory in such a way as to only encode and process highly important information. This ability is important in the western education system, where value is assigned to pieces of knowledge and information, which are then intentionally encoded by students. Away from the education setting, there is also a clear benefit of being able to control and prioritise what information is retained, as it is impossible to maintain and accurately recall all pieces of information you come into contact with. For example, it might be critical for you to remember certain antibiotics you are allergic to, but you will not need to remember all antibiotics you have taken in your lifetime. This shows that in everyday life you have to prioritise and assign value to certain pieces of information over others.

The controlled and strategic effects of reward on memory are typically investigated in intentional encoding tasks, where participants know they will need to later remember some or all of the information presented. Evidence from the value-directed remembering paradigm shows that people can alter memory performance by modulating intentional encoding in response to value (Castel, Farb, & Craik, 2007; Castel et al., 2013; Cohen et al., 2014; Middlebrooks & Castel, 2018). In the value-directed remembering paradigm participants are given word lists to remember and each word is given a point value. The aim for participants is to score the highest points. Participants are told that they will later be tested and for every correctly recalled word they will win that associated point value. It has been consistently found that high value items are recalled more than low value items in this task (Castel et al., 2007, 2011, 2013; Hennessee, Castel, & Knowlton, 2017; Hennessee, Patterson, Castel, & Knowlton, 2019; Middlebrooks & Castel, 2018).

Several different variations of the value-directed remembering paradigm have been developed to investigate the limits of memory prioritisation due to rewards (Ariel et al., 2015; Castel et al., 2007, 2011, 2013; Cohen et al., 2014; Hennessee et al., 2017, 2019; Robison & Unsworth, 2017). In the adapted but widely used value-directed remembering paradigm, participants take part in stages of word list learning tasks and are given free-recall tests after each learning phase. This phased approach yields a trial-and-error learning effect where participants modify their strategy to selectively attend to high value items and ignore low value items as much as possible. In this paradigm, participants learn from experience of their own memory capacity, and update their strategy quickly to maximise points scored. When free recall tests were intermixed in this way this led to increases in recollection and familiarity-based retrieval for high value items, whereas when learning occurred without intermixed testing there was only an enhancement of recollection. The authors suggested that the enhancement of both recollection and familiarity was a result of participants engaging strategic control processes to maximise point scoring, as well as an automatic dopaminergic influence. The recollective enhancement of high value items only in the non-intermixed test conditions was attributed to engagement of the dopaminergic reward systems automatically enhancing higher value item memory. This evidence thus supports the view that rewards can result in automatic dopaminergic driven effects which can interact with separable controlled strategic processes. This paradigm has raised interest with the ageing literature as a potential method to engage strategic control processes to boost memory for important information, such as medication dosage.

It is suggested that strategic prioritisation of control processes leads to an engagement of deeper, more elaborate and semantically engaged processes of high value information

(Cohen et al., 2017, 2014), as well as an automatic attentional draw higher value, or rewarding items (Anderson, 2016). Increased sensitivity to reward value (strategically remembering the highest values only, even if this means less words are recalled) was associated to activity in semantic processing regions (left inferior frontal gyrus and left posterior lateral temporal cortex) for high value words compared to low value words (Cohen et al., 2014). This supports the conclusions that deep, elaborate processing is engaged, without instruction, to selectively use memory capacity and to maximise the points scored. However, when explicit strategies are given to aid encoding, value no longer has such an influence and appears to no longer be a strong modulator of memory performance (Hennessee et al., 2019). This provides more evidence in support of rewards interacting with control processes and strategic encoding prioritisation. When strategies are instructed, these are then applied to encoding across the board and therefore strategies cannot be used only to boost rewarding or important memory.

Overall, it is apparent that external rewards can boost both intentional and incidental encoding, however it is not known from prior literature how external rewards affect incidental encoding during recognition attempts, as addressed in the current thesis.

1.5.5 Rewards & healthy ageing

Because of the potential to boost memory ability, there has been an increased interest among researchers in how reward processing interacts with episodic memory in older age. Awarding rewards could be a helpful method for reducing memory decline seen in healthy ageing. However, factors that have been shown in younger adults to enhance learning, such as external rewards, may not affect older adults in the same way. For example, older adults have been reported to be motivated by internal driving factors such as curiosity and internal achievement

(McGillivray, Murayama, & Castel, 2015; Sakaki, Yagi, & Murayama, 2018), suggesting that they may be less sensitive to external rewards than younger adults.

However, the evidence on age differences in reward sensitivity and effects on memory is very mixed. Some findings have indicated that reward effects on memory are reduced or absent in older adults (Geddes, Mattfeld, Angeles, Keshavan, & Gabrieli, 2018), consistent with evidence for age-related impairments in dopaminergic neural systems such as reductions in physiologically dopamine are linked with reduced reward processing (Bäckman, Lindenberger, Li & Nyberg, 2010). However, other findings suggest that reward effects on incidental encoding may be intact in older adults (Mather & Schoeke, 2011) pointing towards a more complicated relationship between ageing, reward processing, and different types of memory processes and stages. This view is supported by other studies that have found mixed evidence for intact versus reduced reward effects on memory in older age (e.g. Spaniol, Schain, & Bowen, 2014; Geddes, Mattfeld, De los Angeles, Keshavan, & Gabrieli, 2018). For example, external rewards increase the use of strategic memory processes in younger samples who will prioritise encoding of rewarded information at the expense of non-rewarded information (Cohen, Rissman, Suthana, Castel, & Knowlton, 2014), with some evidence suggesting older adults will also strategically prioritise encoding in this way (Castel, Farb, & Craik, 2007; Castel et al., 2013; Cohen, Rissman, Suthana, Castel & Knowlton, 2016). Older adults have been seen to selectively attend to high value items, but they did still show deficits in source information relating to high value items and instead relied on gist based memory (Castel et al., 2007). Interestingly, Eich, Stern, & Metcalfe (2013) found no hypercorrection effects in a healthy ageing population. The authors found that older adults had a good level of general knowledge overall, and also corrected errors successfully, but this was not linked to the initial confidence of the error. It was suggested that this result might be related to the way in which older adults assign confidence to their judgments, or that there might have been less of a surprising effect of the high confident errors in older than younger adults.

Thus, the literature on reward-related effects on memory in older age is somewhat conflicting, which may relate to different studies measuring reward effects on memory that depend on relatively automatic dopaminergic brain systems (eg. Wittmann et al., 2005), versus strategic recruitment of cognitive control processes (eg. Cohen, Cheng, Paller, & Reber, 2019). Current research has only begun to assess the relationship between reward-induced motivation and strategic control processes in terms of their effects on memory (e.g. Cohen, Cheng, Paller & Reber, 2019), and how that relationship may differ across age groups, memory stages (e.g. encoding vs. retrieval), and memory processes (i.e. different types of encoding or retrieval operations, such as intentional vs. incidental encoding/retrieval processes).Hence, these issues were investigated in some of the empirical work in this thesis.

1.6 Overview of thesis

This thesis investigates behavioural and ERP measures of incidental encoding during recognition attempts. Specifically, I investigate different potential new factors that could moderate such encoding, given that these factors are known influences of episodic memory more broadly. As reviewed in this chapter, I focused on effects of healthy ageing and external rewards, and also investigated how those factors interact with retrieval processes that are engaged during recognition attempts.

During retrieval, we may engage strategic processes relatively consistently throughout a situation, such as reinstating encoding strategies or prioritising important information throughout a memory test, yet there are also automatic and controlled memory processes that are elicited by items. These item-specific retrieval processes may be engaged as a result of item properties, such as the degree of familiarity or novelty processing elicited by an item, or may fluctuate depending on task engagement, such as how much attention we allocate to items. Item-specific retrieval processes are often considered in relation to whether or not retrieval is successful, however have not been widely examined in relation to the encoding of novel information during retrieval attempts. I have previously reviewed source-constrained retrieval as one example of strategic retrieval processing that may be quite consistently applied throughout a task and may boost encoding, but post-retrieval monitoring is another controlled process that may be elicited by certain items, that might also be important for learning during retrieval. Likewise, people's subjective experience of memory and how that experience is used during decision making may affect whether information in the retrieval environment becomes encoded. Little is known about the impact of subjective recognition experience on incidental encoding, but insights about prediction errors from the semantic memory field (Butterfield & Metcalfe, 2006), as well as the reward literature (Rouhani et al., 2018), have suggested that these processes may relate to memory updating, suggesting they may also be important for incidental encoding of novel information that is encountered in a retrieval context. Likewise, there have been suggestions that confidence judgments can influence memory modification and updating (Finn, 2017). However, to my knowledge no studies have examined the influence of subjective experience of retrieval and recognition decision making on incidental encoding, as investigated in this thesis.

In Chapter 2, I present my first experiment that aimed to replicate the Jacoby, Shimizu, Velanova, et al., (2005) study that is the only study that has explored the effects of healthy ageing on source constrained retrieval. Experiment 1 expanded on this study by modifying the memory-for-foils paradigm to explore the effects of external rewards on incidental foil encoding, as well as any interaction of external rewards with healthy ageing. Experiment 2, in

Chapter 3, was designed to more directly test the effects of external rewards on incidental encoding rates during recognition by modifying the memory-for-foils paradigm to incorporate aspects of the value-directed retrieval paradigm. As this research was very exploratory, this study did not examine the effect on healthy ageing. Experiment 3, presented in Chapter 4, continued to build on the effect of external rewards on younger adults from Experiment 1, whilst also trying to increase the salience of external rewards. This experiment continued to use the memory-for-foils paradigm, modifying the task used in Experiment 1 by incorporating separate measures of familiarity and recollection (using Remember/Know judgements) and also confidence ratings and feedback. This chapter therefore investigated effects of both external rewards and measures of item-specific recognition processes on incidental encoding during recognition. Experiment 4, presented in Chapter 5, directly examined hypotheses relating to the recognition processes driving differences in incidental encoding rates suggested from Experiment 3. Here I used ERPs to examine the influence of familiarity and novelty processing combined with behavioural measures of confidence, response times and source judgements to examine the influence of retrieval and decision processes on incidental encoding rates. Here I also expanded this topic to examine the effects of healthy ageing on these processes. Finally, in Chapter 6 I provide a general discussion of the key findings and implications of the empirical chapters, together with limitations and future suggestions based on this body of work.

Chapter 2: Source-constrained retrieval in older adults; LOP and Reward manipulations

2.1 Overview

In Chapter 1, I reviewed the previous literature examining encoding of novel items during retrieval, why this occurs, what factors are known to affect this process and suggested some other potential factors that could affect this process.

In this chapter, I present one experiment that aims to replicate and extend on the previous literature. The experiment is based on the source-constrained retrieval literature (Alban & Kelley, 2012; Danckert et al., 2011) and the memory-for-foils paradigm (Jacoby, Shimizu, Velanova, et al., 2005). I aimed to replicate these findings by comparing an older adult sample with a sample of young adults as to my knowledge, there is limited evidence on how older adults reinstate strategies and limited verification of deficits in strategic source-constrained retrieval in older adults. In the current experiment, I also extended on the previous literature to explore a novel factor of external reward. I examine how external rewards affect intentional encoding with the aim to alter retrieval processes and incidental encoding rates, and if this interacts with age.

2.2 Experiment 1

The ability to control retrieval is highly advantageous to the memory system. Young people with intact control processes often self-initiate recollection to enhance the accuracy of their recognition decisions, for example by strategically orienting their memory search towards a particular context (Jacoby, Shimizu, Daniels, & Rhodes, 2005; Mecklinger, 2010; Rugg & Wilding, 2000). The constraining of memory relies heavily on effortful, cognitive control processes, yet it has been shown that older adults are less likely to spontaneously use retrieval strategies in memory tasks (Cohn, Emrich, & Moscovitch, 2008; Kirchhoff, Anderson, Barch, & Jacoby, 2012) and show reductions in cognitive control processes more generally (Luo & Craik, 2008). Yet, only one prior study, to my knowledge, has investigated the effects of ageing on source-constrained retrieval (Jacoby, Shimizu, Velanova, et al., 2005). In that study, Jacoby, Shimizu, Velanova, et al., found a deficit in spontaneous source-constrained retrieval (SCR) in older adults. That is, older adults did not show any difference in subsequent recognition accuracy of foils that had previously been shown in the deep versus shallow recognition tests, whereas a young group of participants showed the standard pattern of enhanced memory for foils from the deep test compared to the shallow test. This experiment attempted to replicate and extend on prior findings of age-related deficits in spontaneous use of controlled retrieval processes (Jacoby, Shimizu, Velanova, et al., 2005), by also investigating if such deficits were accompanied by age-related reductions in reward effects on memory.

Increasing the use of self-initiated strategies in both younger and older samples attracts a lot of interest as a method to memory performance. External rewards have been seen to increase the use of strategic memory processes in younger samples who will prioritise encoding of rewarded information at the expense of non-rewarded information (Cohen, Rissman, Suthana, Castel, & Knowlton, 2014), with some evidence suggesting older adults will also strategically prioritise encoding in this way (Castel, Farb, & Craik, 2007; Castel et al., 2013; Cohen, Rissman, Suthana, Castel & Knowlton, 2016). However, to my knowledge, the effects of reward on encoding have not been investigated when encoding occurs during retrieval attempts, as in the memory-for-foils paradigm. Therefore, this paradigm enables the examination of how rewards could moderate and boost strategic engagement, and therefore incidental encoding.

This first experiment therefore tested whether rewards might enhance incidental foil encoding in similar ways as other factors, such as semantic processing, imagery (Danckert et al., 2011), self-referential thinking (Bergström, et al., 2015), survival processing (Nairne, Pandeirada, VanArsdall, & Blunt, 2015) and full attention (Dudukovic et al., 2009). Reward-induced enhancement of incidental foil encoding could potentially be mediated by automatic, dopaminergic mechanisms (Bäckman, Lindenberger, Li, & Nyberg, 2010), or through the recruitment of strategic memory processes during the initial recognition tests. That is, I predicted that the prospect of external rewards during a recognition test might encourage participants to use cognitive-control dependent strategies such as careful retrieval monitoring or effortful attempts at recollection (Halsband et al., 2012), and such strategies would likely involve enhanced attention and "deeper" processing of foils when compared to processing of foils in a non-rewarded test. Therefore, increased use of retrieval strategies in response to rewards might facilitate incidental foil encoding in younger adults. If strategic retrieval processes are impaired in older age, older adults might show reduced reward effects on incidental foil encoding.

The current study therefore directly compared both source-constrained retrieval effects and reward effects on incidental foil encoding across younger and older adults in the memory-for-foils paradigm. The experiment was conducted with one older group (mean age 73 years old) and one younger group (mean age 20 years old), and included two different manipulations of retrieval processing during recognition tests. The LOP phase was closely based on previous research (e.g. Jacoby, Shimizu, Velanova, et al., 2005; Vogelsang et al., 2016; 2018), and included two initial recognition tests that varied whether foils were shown either together with old items that had been deeply/semantically encoded or shallowly/perceptually encoded, in order to manipulate source-constrained retrieval processing. An additional external reward phase was also added where I conducted two initial recognition tests that varied whether recognition performance was rewarded or not rewarded, in order to manipulate reward-related motivation. Subsequently, recognition memory for all foils from the four prior tests (deep/shallow/rewarded/not rewarded) was tested in a surprise final test.

I predicted that in line with prior findings (e.g. Jacoby, Shimizu, Daniels, et al., 2005; Vogelsang et al., 2016; 2018), final recognition would be enhanced for foils previously shown in the deep compared to the shallow test in the younger group, however there would be no difference in final recognition of foils previously shown in the deep vs. shallow test in the older group. I also expected that surprise test recognition would be enhanced for foils from the rewarded compared to non-rewarded test in young adults, in line with prior evidence for reward-related enhancement of incidental encoding (eg. Mather & Schoeke, 2011; Wittmann et al., 2005). Based on evidence that reward sensitivity (Geddes et al., 2018) and spontaneous use of strategic memory processes (Kirchhoff et al., 2012) it was predicted that a smaller difference would be seen between the reward and no reward foils in the older group compared to the younger group. In contrast, if reward-related motivational influences on memory are intact in older age (Mather & Schoeke, 2011), then both older and younger adults would be expected to show similar reward-related enhancements of foil recognition.

2.3 Methods

2.3.1 Participants

I aimed to recruit 48 participants per group in order to achieve >0.9 power to detect an effect size of Cohen's *d*=0.7 at two-tailed α=.05, and to fully rotate and cross all counterbalancing factors (task orders, stimuli assignment to conditions, etc.). This estimated effect size was based on previous research with a similar design and using similar stimuli, which found large effects of source-constrained retrieval on foil recognition in young adults (Cohen's *d*=0.75 and *d*=0.89 for recognition accuracy differences between Deep vs. Shallow foils on the surprise test, in

Vogelsang et al., 2016; 2018). The final sample was reduced to 41 per group partly due to practical recruitment constraints and partly due to participant exclusions, resulting in >0.9 power to detect an effect size of Cohen's *d*=0.75 at two-tailed α=.05. Specifically, an additional nine young adults were replaced or excluded; three due to technical errors and six due to scoring below threshold on the Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005). An additional six older adults were also excluded; three due to technical errors, two due to MoCA scores below threshold, and one due to a failure to follow task instructions. Six of the younger adults and three of the older adults in the final sample reported that they anticipated a final test of some kind, but these participants were not excluded as the pattern of results was not affected by their removal. Thus, the presented results are based on data from 41 young adults (*M* age = 19.6 years, *SD*=1.1, range = 18-24) and 41 older adults (*M* age = 72.6 years, *SD*=5.2, range= 64-85; based on *N*=40 due to one participant missing demographic data). Note however that the sample size is still larger than that used by Jacoby et al. (2005) in their most similar experiment (Experiment 1), which included 32 younger (mean age 19.6, range 18-26) and 32 older (mean age 75.8, range 61-87) adults in a memory-for-foils test that manipulated LOP as a between-subjects factor (thus testing 16 young and 16 older participants in each condition).

Both the age groups had scores within the normal range on the Montreal Cognitive Assessment (Nasreddine et al., 2005; MoCA; cut off >= 25[1]; young adults *M* = 28.0, *SD*= 1.3; older adults *M* = 27.7, *SD*= 1.8). Older adults were highly educated (mean number of years in education = 17.7, *SD*= 4.4; based on *N*=40 due to one participant missing demographic data) and young adults were all currently in higher education (mean number of years in education = 14.5, *SD*= 1.4). The young adults were recruited from the University of Kent Psychology

---

[1] The MoCA threshold of >=25 was used instead of the standard 26, in order to be less conservative with regards to inclusions. Since the main results showed a lack of memory difference between the age groups, we did not want to restrict group differences by being overly conservative. For the record, all key conclusions from analyses were the same when a MoCA cut off of >=26 was used.

Department research participation scheme, and the older adults were recruited from the Canterbury District of the University of the Third Age. All participants were native English speakers, had normal or corrected to normal vision (colour-blindness was not tested and only self-reported), reported not having depression, anxiety or dyslexia, and were not taking any psychoactive medication. Participants could have been left or right handed. All participants provided full written informed consent and 'won' the same monetary reward of £5 as part of the experimental manipulation. Young participants also received course credits in addition to the reward. The procedures were approved by the University of Kent, School of Psychology Ethics Board. All testing was conducted in a face-to-face manner.

2.3.2 Materials

The experiment was programmed and run in PsychoPy (Peirce, 2007) on a tablet computer. In total 288 English words from the MRC psycholinguistic database (Wilson, 1988) were used as stimuli (3-11 characters, Kucera-Francis frequency range 0-591), drawn from a larger set used by Vogelsang et al. (2016, 2018). The words were randomly split into twelve lists of 24 words each (see Appendix 1.2 for full characteristics of the word list break down), with assignment to experimental conditions counterbalanced across participants. Counterbalancing of word list was also factored into counterbalancing of the manipulation (LOP or Motivation) and manipulation level presentation (ie; Deep-Shallow or Shallow-Deep). This yielded a counterbalancing rotation of 24 participants per age group. No counterbalancing of handedness was included to ensure instructions were easy to follow for the older adult experimenters. I also decided that given that behavioural results of accuracy were the primary measure counterbalancing was not necessary, as differences in RT between 'Old' and 'New' responses was not of interest here and the interaction with manipulation factors can still be examined here.

A set of 160 emotionally neutral photographs from the IAPS database (Lang, Bradley, & Cuthbert, 2005) were used in a visual search filler task. Each image was available for use twice (360 images). This created a set of 8 lists with 40 images in each, with each list having different target images. A random selection of 10 images from each list had a white triangle superimposed on top of the image somewhere, this was the target image. The triangle varied in location and size, but was clearly not part of the image itself. Participants were instructed not to focus on the image but to search for the white triangle.

2.3.3 Design and Procedure

The study was conducted as a "citizen science" project, where I collaborated with the University of the Third Age (U3A) research group members who were older adults themselves and acted as advisors and experimenters. The aim of this collaboration was to mitigate potential confounding effects of experimenter and participant expectations for stereotyped age differences in memory (McDaniel, Einstein, & Jacoby, 2008). The U3A research group was consulted during the study design to make the task more suitable for older adults, and were extensively trained in order to independently conduct data collection, with assistance from *LS*. Thus, I took great care to ensure that all tasks were suitable for both age groups based on consultation with the U3A research group members who were of similar age as the older sample. As a result of this consultation, all instructions were given both verbally and in writing, multiple breaks were provided at specific points during the task, and computerised response requirements were simplified and consistent across all tasks. All materials, task designs and instructions were identical across both groups, and instructions were delivered without mentioning that memory performance would be compared across age groups, to avoid expectancy effects. The experimenters' ages were consistent with the age groups that they tested; the older adults were tested by an older adult experimenter from the U3A research group (together with *LS*), whereas the younger adults were tested by a younger adult experimenter (either *LS* or an undergraduate student assistant).

2.3.3.1 Main encoding and recognition tasks

The memory-for-foils procedure (Jacoby, Shimizu, Daniels, et al., 2005) was modified to incorporate two different encoding manipulations, Level of Processing (LOP) and Reward, which were presented to the participants in separate blocks and were followed by a joint surprise test for all foils. The key variable that changed across these blocks was what type of encoding task participants were asked to complete. As seen in Figure 2.1 (schematic of the task order), within each manipulation block, participants first completed two encoding phases (either Deep and Shallow, or Rewarded and Non-Rewarded), then went on to complete two separate recognition tests for old items drawn from each of the encoding tasks, together with randomly intermixed new foils. The recognition tests were always conducted in the same order as the encoding tasks, and participants were told which encoding task the old items had previously been shown in. The manipulation order (LOP/Reward or Reward/LOP) was counterbalanced across participants, as was the specific condition order within manipulation block (i.e. Deep/Shallow or Shallow/Deep within the LOP block, and Rewarded/Non-Rewarded or Non-Rewarded/Rewarded within the Reward block). So, for example, a specific participant would have started with a Deep encoding task followed by a Shallow encoding task, then had their recognition of Deeply encoded old items tested intermixed with foils, followed by a recognition test for Shallowly encoded old items intermixed with foils. Next, they would have completed the Rewarded encoding task, followed by the Non-Rewarded encoding task. Then, they would have competed a recognition test for the Rewarded old items intermixed with foils, then another recognition test for the Non-Rewarded old items intermixed with foils. Other participants completed these blocks and conditions in counterbalanced orders so that there were no order confounds in the design, with the constraint that LOP conditions were always presented together in one block and Reward conditions in the other, and the initial recognition tests were always conducted in the same order as the encoding tasks, so that the time delay between encoding and recognition was always constant. Finally, all participants completed a Surprise recognition test at the end where foils from all four recognition tests were randomly intermixed with completely new words.

The LOP phase was very similar to prior research with the memory-for-foils paradigm (e.g. Jacoby, Shimizu, Velanova, et al., 2005). In this phase, participants encoded two different lists of 24 words during separate Deep vs. Shallow encoding tasks (order counterbalanced). In the 'Deep' encoding task, they were asked to judge whether each word was pleasant, whereas in the 'Shallow' encoding task, they were asked to judge whether each word contained an 'O' or a 'U'. Words were presented for 3 seconds and participants pressed 'Z' or 'M' keys on the keyboard within this time to indicate 'Yes' or 'No' to the relevant question. Each word was preceded by a 600ms fixation cross.

Participants then took part in two recognition tests (the Test 1 stage), in the same order as the encoding tasks (i.e. if the Deep encoding task was conducted before the Shallow encoding task, then Deep old items were tested before Shallow old items, and vice versa). Each test contained 24 old target words and 24 new foil words. As in previous research (Jacoby, Shimizu, Daniels, et al., 2005; Jacoby, Shimizu, Velanova, et al., 2005; Vogelsang et al., 2016; 2018) participants were explicitly told that the 'old' words in these separate recognition tests were drawn from only one of the earlier judgment tasks (Deep or Shallow) and that they were not intermixed. Their task was to say whether they recognised each word as having been shown in the relevant earlier judgement task. These instructions thus aimed to hint to participants that it could be useful for them to strategically reinstate the processing mode they had used during the corresponding study task; however, participants were not explicitly asked to do so. Therefore, any reinstatement of Deep vs. Shallow processing during the recognition test is considered to have been spontaneous. Participants again used the ‘Z’ or ‘M’ keys to respond ‘Yes’ or ‘No’, this time to the question: ‘Do you recognise the word?’. Each word was preceded by a 600ms fixation cross, and remained on the screen until participants gave a response. Participants were asked to respond quickly and accurately.

In the External Reward phase, participants encoded two different lists of 24 words during separate Rewarded vs. Non-rewarded encoding tasks (order counterbalanced). They were told that for one list (the Reward condition), they would receive a monetary reward for correctly recognising those words later (10p per word, max £5), whereas there would be no monetary reward for the other list. As in the LOP encoding tasks, all words were presented for 3 seconds, preceded by fixation cross for 600 milliseconds, but no judgments or responses were required during the Reward phase encoding tasks. Participants studied both lists before moving onto the corresponding recognition test phases. As in the LOP phase, word lists were tested in two separate “Test 1” tests by intermixing old words from the encoding tasks with new foils (24 old and 24 new words in each test), in the order of study presentation (i.e. if the Rewarded encoding task was conducted before the Non-rewarded encoding task, then Rewarded old items were tested before Non-rewarded old items, and vice versa). Participants were once again explicitly told which test corresponded to which study list so that they knew whether they would be receiving rewards or not, based on their recognition performance. Participants were informed that they would be rewarded for correct responses to both old and new words. Rewards were balanced across old and new items to avoid inducing response biases, which could have emerged if only correct “old” responses were rewarded (such a design could encourage participants to adopt a “lax” response bias and respond ‘old’ to the majority of items in an effort to maximise their winnings). For ethical reasons, but unknown to participants, at the end of the experiment all participants received the full £5 win. As in the LOP recognition tests, each word was preceded by a 600ms fixation cross, and remained on the screen until participants gave a response. Participants were asked to respond quickly and accurately.

After all the LOP and Reward encoding phases and initial recognition test phases had been completed, participants received a surprise final recognition task for all the 96 foil words that had appeared in the four tests (24 Deep, 24 Shallow, 24 Rewarded, 24 Non-Rewarded foils), intermixed with an equal number of new words (96 new words). Participants were instructed that in this test, words that had been presented previously in any of the blocks could appear, intermixed with completely new words that had never been presented during the study. It was ensured that participants understood that this included words from all previous phases, including the 'foil' words that had been presented in the previous test phases, and that their task was to say whether they recognised each word from any of the previous phases. As in previous tests, participants again used the 'Z' or 'M' keys to respond 'Yes' or 'No' to the question: 'Do you recognise the word?' as quickly and accurately as possible and the task was self-paced, with each word preceded by a 600ms fixation cross. (See Fig 2.2 for a schematic of the trial timing across all experimental phases).

2.3.3.2 Visual Search filler task

A visual search task was used as a filler task to prevent rehearsal between each experimental task (eight times in total). Participants were presented with a series of photographs and were asked to search for a white triangle that had been superimposed onto a subset of photographs (See Fig 2.3). As in previous tasks, participants again used the 'Z' or 'M' keys to respond 'Yes' or 'No', this time to the question, 'Is there a white triangle on the image?'. This filler task always lasted 1 minute for all blocks and all participants, regardless of the number of trials completed (i.e. faster participants were given more trials, and slower participants were given fewer trials, in order to ensure a constant delay between tasks).

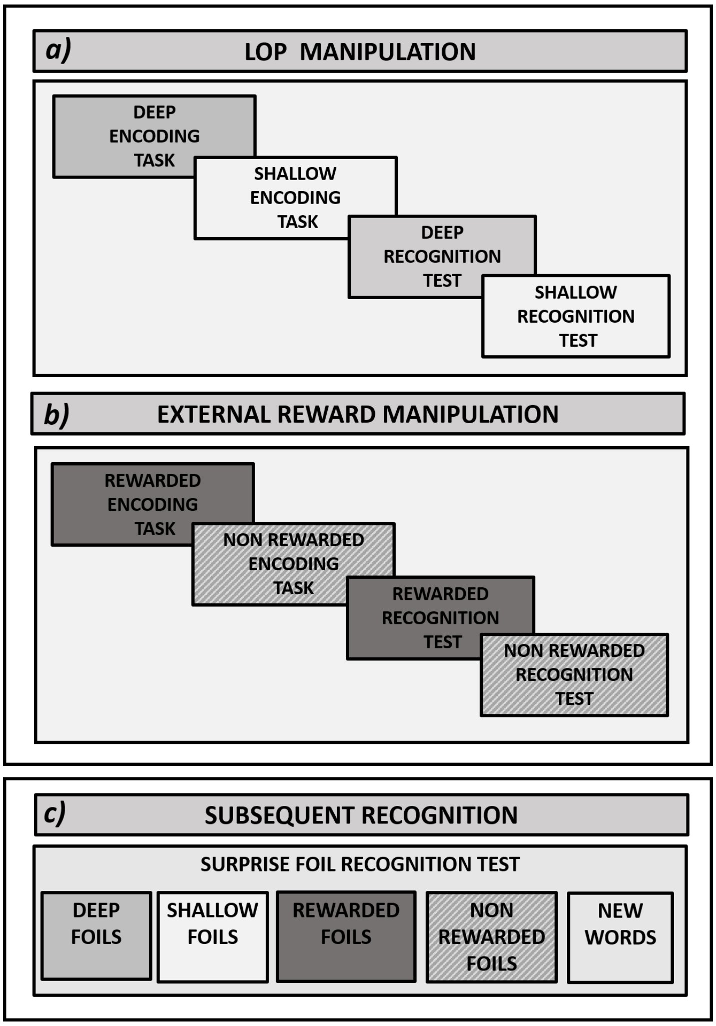


Figure 2.1. Schematic of Experiment 1 procedures. Every participant took part in both the LOP (*a*) and the External Reward manipulation (*b*) before taking part in the subsequent recognition phases which was a surprise recognition test (*c*) containing all the four foil types and new items. The order of LOP (*a*) External Reward (*b*) manipulations were counterbalanced across participants, as was the order of each manipulation level (LOP: deep-shallow order and External Reward: rewarded-non rewarded order).

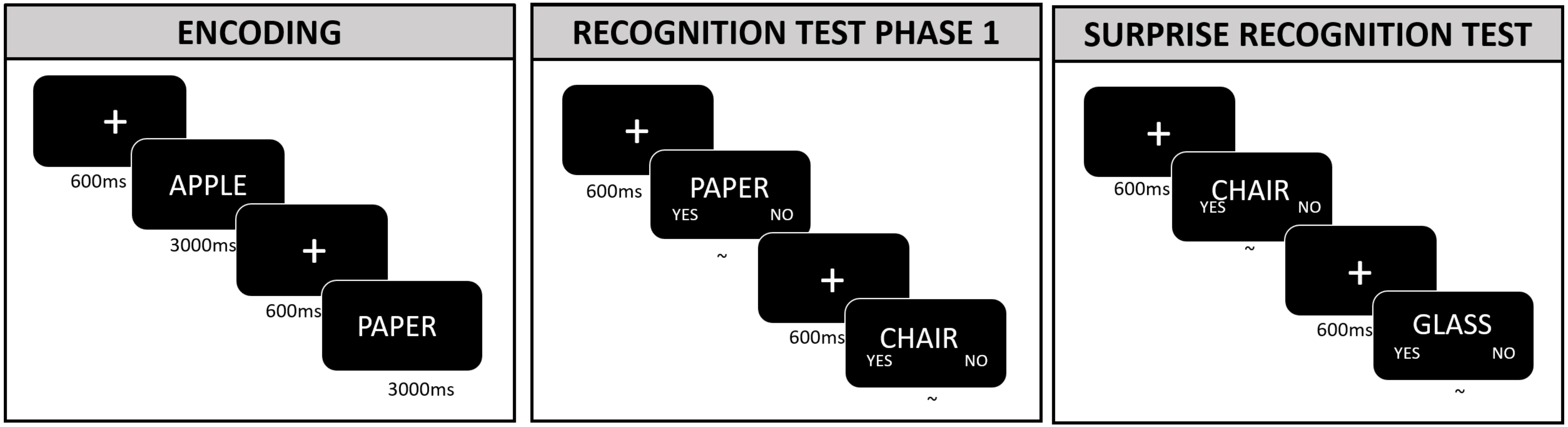


Figure 2.2. Schematic of Experiment 1 trial timings, all timings are in milliseconds (*ms*) unless self-paced task which is indicated by the tilde symbol (~). Trial timings were consistent across all manipulations.

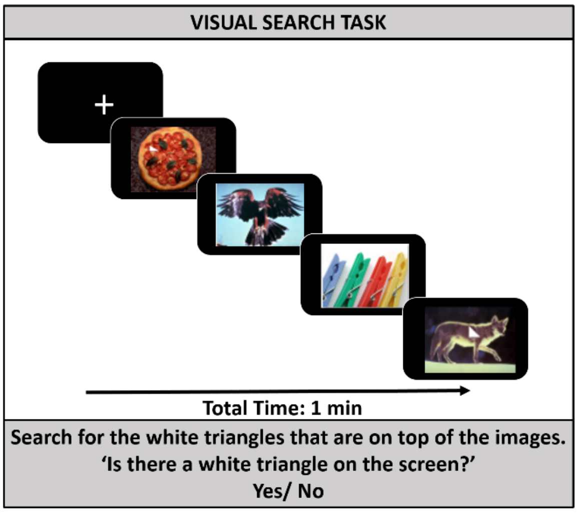


Figure 2.3. Schematic of Visual Search Task procedures. Every participant took part in the visual search filler task 8 times in between any other task stage. Each task lasted 1 minute in total regardless of trials completed and each trial was self-paced.

2.3.3.3 Post- Experiment tasks and Debrief

After the participants had completed the experimental tasks, they completed a short questionnaire, assessing if they had anticipated the final foil test as well as questions about effort, motivation and strategies used during the reward phases (see Table 2.5 for the abbreviated questions asked). They also completed the MoCA (Nasreddine et al., 2005), and were given a verbal as well as a written debrief alongside their reward phase earnings of £5.00.

2.3.4 Data Analysis

For the Recognition Test 1 stages, the proportion of correct responses for each test block (Deep, Shallow, Reward, No Reward) was calculated separately for Old words (i.e. hit rates) and New words (i.e. correct rejection rates) within each test block. Raw accuracy scores were used to examine if there were differences in specific responses between groups to old or new words, which may not be observed by looking at discrimination and bias alone.

Mean reaction times (RTs) were also calculated for Old and New words split by test block type. Reaction times were primarily used descriptively to characterise performance and were not analysed to test key hypotheses (which focused on accuracy). Although RTs may be more reliably assessed using median rather than mean values, I extracted the means to provide a comparable measure to prior literature (Bergström et al., 2015; Jacoby, Shimizu, Daniels, et al., 2005; Vogelsang et al., 2016) which also used mean RT scores. Individual mean RT scores are shown where relevant to show the distribution of RT responses across the sample.

The raw accuracy scores above were used to derive estimates of discrimination ability and response biases during each recognition test, in order to provide more meaningful measures of recognition performance for testing potential differences between the conditions and age groups. *Pr* was calculated to measure participants' ability to discriminate between old and new items regardless of response biases. *Pr* represents how much more likely participants are to accurately identify previously seen items as old, compared to mistakenly "false alarm" by identifying new items as old (*Pr* = hits- false alarms; see Snodgrass & Corwin, 1988). The *Br* measure of response bias was also calculated (*Br* = false alarm / (1- *Pr*); Snodgrass & Corwin, 1988) to estimate participants' general tendency to respond 'old' versus 'new' when uncertain. *Br* values can fall between 0-1, where values around 0.5 indicate a neutral response bias (equally likely to guess old or new), values >0.5 indicate a "liberal" response bias (guessing "old" more often than "new") and values <0.5 indicate a "conservative" response bias (guessing "new" more often than "old").

On the Surprise foil test, proportion accurate responses were calculated separately for each foil condition (Deep foils, Shallow foils, Reward foils, Standard foils), and for New items (resulting in hit rates for foils since these items are "old" on the final test, and correct Rejection rates for New items). Mean RTs were also calculated for these conditions. As there was only one set of New items and all Foils were intermixed, one measure of Surprise test response bias

(*Br*) was calculated for each participant, using a summary measure of *Pr* that was not split according to foil condition (*Br* = false alarm / (1- summary *Pr*). Condition-specific *Pr* scores were then calculated for each foil type by subtracting the common false alarm rate from the separate foil hit rates.

Pr was used as the key measure of recognition discrimination ability for inferential statistical analyses of both Test 1 and Surprise test data, in order to make the results comparable across phases and to compensate for individual differences in response biases. The *Br* measure was also included as a key measure to investigate potential age and condition differences in response biases. Raw proportion accurate measures (hit and correct rejection rates) were also analysed, but these measures are only complementary to the key *Pr* and *Br* measures. Likewise, differences in RTs across conditions were also analysed. All analyses compared measures across conditions and groups within the LOP and Reward manipulation types, but did not compare measures across LOP and Reward manipulations, since such comparisons would not be meaningful.

The key analyses used frequentist inferential statistical tests from the General Linear Model (ANOVA, t-tests). Effect sizes were estimated using partial eta squared (ANOVA) or Cohen's *d* (t-tests). Cohen's *d* for both paired and independent t-tests was calculated as the difference between means divided by the pooled standard deviation to avoid inflating effect size estimates for paired t-tests (Dunlap, Cortina, Vaslow, & Burke, 1996). Spearman Rho correlations are used when non-parametric Pearson's correlations would not be suitable, for example with ranked variables. Frequentist statistical tests of key hypotheses were supplemented with Bayesian statistics in order to assess if evidence supported the null hypothesis, which is not possible to determine with frequentist statistics alone. Therefore, Bayes factors (BF_{10}) were calculated to estimate the relative support for the alternative (H_1) versus the null hypothesis (H_0). The Bayes Factor is a ratio that contrasts the likelihood that the data would occur under the alternative (H_1) versus null (H_0) hypotheses, with values over 1 indicating support for H_1 and values below 1/3 indicating support for the H_0. Values close to 1 are only considered weakly/anecdotally supportive of one hypothesis over the other, whereas $BF_{10} > 3$ are typically interpreted as substantial evidence in support of H^1 over H^0, and $BF_{10} < 1/3$ are interpreted as substantial evidence in support of H^0 over H^1 (see Wagenmakers, Wetzels, Borshboom & Van der Mass, 2011). All Bayes factors were calculated with two-tailed tests, thus the alternative hypothesis was a statistical difference between groups/association between variables in either direction, and the null hypothesis was no difference between groups/no association between variables. Bayes Factors were calculated using JASP (JASP Team, 2017) with default priors (a Cauchy distribution with center = 0, r = 0.707).

2.4 Results

Both group level and individual participant's data from all recognition tests in all phases can be found in an anonymised format on the Open Science Framework website along with other supplemental materials and additional analyses, such as Bayesian analysis output including robustness checks (https://osf.io/aejxm/?view_only=4ada61ec992b4f7288d0ded2697139f8). Some supplementary analysis investigating manipulation order can be found in Appendix 1.3.

2.4.1 Recognition Test 1

2.4.1.1 LOP Manipulation

Performance on the first recognition tests from the LOP phase is shown in Table 2.1. As can be seen here, both age groups were highly accurate on the Deep recognition test, and less accurate on the Shallow recognition test, as expected. The older adults responded slower than younger adults, but discrimination and response biases were similar across groups.

Table 2.1. *Performance on the first recognition tests in the LOP phase. The table shows proportion accurate responses, Discrimination (Pr), Response bias (Br) and mean reaction times, and split by age group and type of preceding encoding task.*

Age Group	Recognition accuracy				Discrimination (Pr)		Response Bias (Br)		Reaction Times (ms)			
	Old Words		New Words						Old Words		New Words	
	Deep	Shallow	Deep	Shallow	Deep	Shallow	Deep	Shallow	Deep	Shallow	Deep	Shallow
Younger Adults	.94 (.07)	.57 (.19)	.92 (.13)	.83 (.14)	.83 (.14)	.38 (.20)	.49 (.27)	.31 (.19)	871 (146)	1131 (317)	984 (185)	1110 (264)
Older Adults	.93 (.09)	.55 (.18)	.92 (.11)	.82 (.14)	.81 (.14)	.36 (.18)	.51 (.24)	.29 (.17)	1165 (359)	1567 (413)	1399 (437)	1454 (421)

Note: Values shown are Means (SD).

The results of a mixed ANOVA assessing effects of the within-subjects factors LOP (Deep vs. Shallow test) and Word Type (Old vs. New word) and the between-subjects factor Age Group (Young vs. Old) on proportion accurate responses is shown in Table 2.2. This analysis showed significant main effects of LOP condition (*EMM;* Deep= 0.93, *SE*=0.01; Shallow= 0.69, *SE*= 0.01) and Word Type (*EMM*; Old word = 0.75, *SE*= 0.01; New word= 0.87, *SE*=0.01) on recognition accuracy, due to higher accuracy for the Deep than Shallow test, and higher accuracy for New than Old Words, overall. There was no main effect of Age Group (*EMM*; Older = 0.81, *SE*=0.01; Younger =0.81, *SE*=0.01). There was a significant interaction between LOP condition and Word Type, caused by a larger accuracy difference between Deep and Shallow conditions for the Old Words (*t*(81)=17.83, *p*<.001, *d*= 2.62; BF$_{10}$ =3.500^{e+26}) compared to the New Words (*t(*81)=6.42, *p*<.001, *d*= 0.73; BF$_{10}$ =1.313^{e+06}). There was no interaction between LOP condition and Age Group nor between Word Type and Age Group, and the three-way interaction between LOP condition, Word Type and Age Group was far from significant. Thus, Older and Younger adults had very similar performance on the first test in the LOP phase.

Table 2.2. *The results of two mixed ANOVAs, assessing effects of LOP (Deep/Shallow) x Word Type (Old/New) x Age Group (Older/Younger) on proportion accurate responses and reaction times in the Test 1 phase.*

Measure	Effect	*F*	*p*	n_p^2
Accuracy	LOP	**392.09**	**<.001**	**.83**
	Word Type	**44.18**	**<.001**	**.36**
	LOP * Word Type	**101.23**	**<.001**	**.56**
	Age	0.4	.527	<.01
	LOP * Age	0.01	.920	<.01
	Word Type * Age	0.16	.690	<.01
	LOP * Word Type * Age	<.01	.957	<.01
Reaction Times	LOP	**45.15**	**<.001**	**.36**
	Word Type	**7.11**	**.009**	**.08**
	LOP * Word Type	**46.24**	**<.001**	**.37**
	Age	**36.97**	**<.001**	**.32**
	LOP * Age	0.33	.569	<.01
	Word Type * Age	0.14	.709	<.01
	LOP * Word Type * Age	**9.05**	**.004**	**.10**

Note. Significant values below .05 are shown in bold. *N*= 82

A mixed ANOVA with the within subjects factors LOP (Deep vs. Shallow test) and the between subjects factor Age Group (Young vs. Old) showed only a main effect of LOP on *Pr* (*F*(1,80)=389.78*, p*<.001*, n^2p*=.83, *EMM*; Deep = 0.82, *SE*=0.02, Shallow = 0.37, *SE*= .02) with no main effect of Age Group on discrimination (*F*(1,80)= 0.39, *p*=.532, n^2p<.01, *EMM*; Older = 0.58, *SE*=0.02; Younger =0.60, *SE*=0.02), and no interaction between LOP and Age Group (*F*(1,80)=0.02, *p*=.90, n^2p<.01). Thus, both younger and older adults were better at discriminating between old and new words in the Deep compared to Shallow test, with no evidence for a difference between the groups. To confirm that the effects of the LOP manipulation on discrimination performance was indeed equal across Age Groups, a Bayes factor was calculated for an independent t-test that compared the difference in *Pr* between LOP conditions (Deep *Pr* minus Shallow *Pr*) across the two Age Groups. The Bayes factor provided substantial evidence in support for the null hypothesis of no difference in discrimination between the Age Groups (BF_{10}= 1/4.32).

Participants showed a more conservative response bias (i.e. a reduced tendency to guess "old" when uncertain) in the Shallow compared to the Deep test (main effect of LOP: *F*(1,80)=46.90*, p*<.001*, n^2p*=.37, *EMM*; Deep= 0.50, *SE=0.*03, Shallow= 0.30, *SE*= 0.02). There was no main effect of Age Group (*F*(1,80=<.001, *p*=.994, n^2p<.01, *EMM*; Older = 0.40, *SE=0.*03; Younger =0.40, *SE=0.*03) nor interaction with age group (*F*(1,80)=0.43*, p*=.516, n^2p<.01). To confirm that the LOP effect on response bias was equal across Age Groups, a Bayes factor was calculated for an independent t-test that compared the difference in *Br* between LOP conditions (Deep *Br* minus Shallow *Br*) across the two Age Groups. The Bayes factor provided substantial evidence in support for the null hypothesis of no difference in LOP effect on response bias between the Age Groups (BF_{10} = 1/3.61).

In relation to recognition ability, measures of discrimination and response bias on the first tests in the LOP phase did not differ across Age Groups, who both showed increased discrimination and a less conservative response bias on the Deep compared to the Shallow test.

A mixed ANOVA (Table 2.2) with the within subjects factors LOP (Deep vs. Shallow test) and Word Type (Old vs. New word) and the between subjects factor Age Group (Young vs. Old) showed main effects of LOP condition on response times (*EMM;* Deep = 1105, *SE*=30, Shallow = 1315, *SE*= 38), Word Type (*EMM*; Old Word = 1184, *SE*= 31; New Word= 1237, *SE*=34), and Age Group *(EMM;* Older = 1396, *SE*= 44; Younger = 1024, *SE*= 43). Furthermore, there was a three-way interaction between LOP condition, Word Type and Age Group, and a significant interaction between LOP condition and Word Type. There was no interaction between LOP condition and Age Group nor between Word Type and Age Group on response time.

LOP x Word type ANOVAs were conducted separately for each Age group to follow up on the three-way interaction. Within the Older adults, Response Times were faster in the Deep than Shallow condition (*F*(1,40)= 19.11, *p*<.001, n^2p=.32, *EMM*; Deep =1282, *SE*=57, Shallow =1511, *SE*= 62) and there was a trend for faster Response Times to Old words than to New words (*F*(1,40)= 3.32 p=.076, n^2p=.08, *EMM*; Old Word =1366, *SE*=52, New Word =1427, *SE*=60). There was also a highly significant LOP x Word type interaction (*F*(1,40)=31.51, *p*<.001*,* n^2p=.44), which was due to Older adults being significantly faster at responding to Old words in the Deep than Shallow test (*t*(40)=6.60, *p*<.001, *d*=1.04; BF_{10}=208538), but this was not the case for responses to New words (*t(*40)=.91, *p*=.370, *d*=0.22; BF_{10} = 1/ 4.04). The pattern of Response Times across conditions was similar in the Younger adults, with faster Response Times in the Deep than Shallow test (LOP main effect (*F*(1,40)= 31.04, *p*<.001, n^2p=.44, *EMM*, Deep=928, *SE*=22, Shallow=1120, *SE*= 43) and faster Response Times to Old than New words (*F*(1,40)= 4.32 *p*=.044, n^2p=.10, *EMM*, Old Word=1001, *SE*=32, New Word=1047, *SE*=31), as well as a significant LOP x Word type interaction (*F*(1,40)=15.18, *p*<.001*,* n^2p=.28). Younger adults similarly responded to Old words in the Deep test significantly faster than to Old Words in the Shallow test (*t*(40)=6.22, *p*<.001, *d*= 1.05; BF_{10} = 45922.77). However unlike the Older adults, Younger adults also responded to New words faster in the Deep than Shallow test (*t(*40)=3.56 *p*=.001, *d*= 0.55; BF_{10}= 36.19).

Thus, Response Times showed that older adults were generally slower to respond than younger adults (as is typically found in the literature) and also showed some differences in how they responded to new items, compared to the younger group. However, those reaction times differences were not accompanied by any age differences in proportion accurate responses, so cannot be interpreted as evidence of ageing-related memory impairments.

So, to summarise, accuracy, discrimination and response bias on the first tests in the LOP phase did not differ across Age Groups, who both showed increased accuracy and discrimination and a less conservative response bias on the Deep compared to the Shallow test. Both groups also responded faster on the Deep than the Shallow test, but for the Older adults this difference was restricted to Old items. Older adults also showed generally slower Response Times overall than the Younger adults.

2.4.1.2 Reward manipulation

Performance on the first recognition tests from the Reward manipulation phase is shown in Table 2.3.

Table 2.3. *Performance on the first recognition test in the Reward manipulation phase. The table shows proportion accurate responses, Discrimination (Pr), Response bias (Br) and mean reaction times, split by age group and type of preceding encoding task.*

Age Group	Recognition accuracy				Discrimination (Pr)		Response Bias (Br)		Reaction Times (ms)			
	Old Words		New Words						Old Words		New Words	
	Reward	No Reward	Reward	No Reward	Reward	No Reward	Reward	No Reward	Reward	No Reward	Reward	No Reward
Younger Adults	.78 (.21)	.69 (.22)	.86 (.16)	.83 (.16)	.60 (.27)	.50 (.27)	.42 (.23)	.35 (.23)	1014 (239)	1020 (256)	1076 (255)	1005 (163)
Older Adults	.76 (.16)	.79 (.14)	.86 (.11)	.87 (.12)	.60 (.19)	.64 (.21)	.41 (.24)	.39 (.18)	1350 (413)	1348 (372)	1390 (397)	1433 (442)

Note: Values shown are Means (SD).

A mixed ANOVA (Table 2.4) with the within-subjects factors Reward condition (Reward vs. No Reward test) and Word Type (Old vs. New Word) and the between subjects factor Age Group (Young vs. Old) showed no main effect of Reward manipulation on Test 1 accuracy (*EMM*; Reward = 0.81, *SE*=0.01, No Reward = 0.80, *SE*= .01) and no effect of Age group (*EMM*; Older = 0.82, *SE*=0.02; Younger = 0.79, *SE*=.02), but overall higher accuracy for New than Old words (*EMM*; Old words = 0.75, *SE*= 0.02; New words= 0.86, *SE*=0.01. There was also a trend towards a significant three-way interaction between Reward condition, Word Type and Age and a significant interaction between Reward condition and Age Group, but no significant interactions between Reward condition and Word Type, nor between Word Type and Age Group.

Table 2.4. *The results of two Mixed ANOVAs, assessing effects of Reward (Reward/No Reward x Word Type (Old/New) x Age Group (Older/Younger) on proportion accurate responses and reaction times in the Test 1 phase.*

Measure	Effect	*F*	*p*	n_p2
Accuracy	Reward Manipulation	1.65	.203	.02
	Word Type	**21.71**	**<.001**	**.21**
	Reward Manipulation * Word Type	1.42	.237	.02
	Age	2.31	.133	.13
	Reward Manipulation * Age	**7.89**	**.006**	**.09**
	Word Type * Age	0.32	.571	<.01
	Reward Manipulation * Word Type * Age	3.94	.051	.05
Reaction Times	Reward Manipulation	0.07	.786	<.01
	Word Type	2.88	.094	.04
	Reward Manipulation * Word Type	0.19	.666	<.01
	Age	**32.00**	**<.001**	**.29**
	Reward Manipulation * Age	1.37	.246	.02
	Word Type * Age	0.6	.441	<.01
	Reward Manipulation * Word Type * Age	2.79	.099	.03

Note. Significant values below .05 are shown in bold. *N*= 82

Reward x Word Type ANOVAs were conducted separately for each Age Group to follow up on the trend-level three-way interaction. Within the Older adults, there was no main effect of Reward condition ($F(1,40)=1.91$, $p=.175$, $n^2p=.05$, *EMM;* Reward = 0.81, $SE=0.02$, No Reward = 0.83, $SE= 0.02$) but a highly significant effect of Word Type with higher accuracy for New than Old Words ($F(1,40)= 14.36$, $p<.001$, $n^2p=.26$, *EMM;* Old Word = 0.78, $SE= 0.02$; New Word= 0.87, $SE=0.02$), but no interaction between these factors ($F(1,40)=0.35$, $p=.556$, $n^2p<.01$). For Younger adults in contrast, accuracy was significantly higher in the Rewarded than Non-Rewarded test ($F(1,40)=6.02$, $p=.019$, $n^2p=.13$, *EMM*; Reward = 0.81, $SE=0.01$, No Reward = 0.80, $SE= 0.01$) as well as for New compared to Old Words ($F(1,40)=9.64$, $p=003$, $n^2p=.19$, *EMM*; Old Word= 0.88, $SE= .02$; New Word= .95, $SE=.02$), and there was also a significant interaction between these factors ($F(1,40)=4.54$, $p=.039$, $n^2p=.10$). Within the Younger adults, Old Words were more accurately recognised in the Rewarded than Non-Rewarded test ($t(40)=2.78$, $p=.008$, $d= 0.42$; $BF_{10} = 4.81$), but there was no difference in accuracy for the New Words across tests ($t(40)=1.01$ $p=.318$, $d= 0.14$; $BF_{10} = 1/3.68$).

The *Pr* measure of discrimination showed no main effect of Reward condition ($F(1,80)=1.71$, $p=.195$, $n^2p=.02$, *EMM*; Reward=0.60, $SE=0.03$, No Reward=0.57, $SE=0.03$) or Age Group ($F(1,80)= 2.28$, $p=.135$, $n^2p=.03$, *EMM*; Older=0.62, $SE=0.03$; Younger=0.55, $SE=0.03$), but there was a significant interaction between Age Group and Reward condition ($F(1,80)=7.88$, $p=.006$, $n^2p=.09$). The Reward manipulation had no significant effect on Older adults' ability to discriminate between Old and New Words ($t(40)= 1.36$, $p=.182$, $d=0.20$) whereas Reward did lead to higher discrimination performance for Younger adults compared to No Reward in the same group ($t(40)= 2.47$ $p=.018$, $d= 0.41$). To provide complementary evidence that the Reward manipulation did indeed affect discrimination performance differently across Age Groups, a Bayes factor was calculated for an independent t-test that compared the difference in *Pr* between Reward conditions (Reward *Pr* minus No Reward *Pr*) across the two Age Groups. The Bayes factor provided substantial evidence in support for the alternative hypothesis that the effect of Reward on discrimination was different between the Age Groups ($BF_{10} = 6.503$).

The response bias index (*Br*) indicated that both Age Groups responded fairly conservatively when uncertain, with a slightly more conservative response bias in the Non-Rewarded test than the Rewarded test (*F*(1,80)=4.38*, p*=.04*,* n^2p=.05, *EMM*; Reward = 0.41, *SE*=0.03, No Reward =0.37, *SE*=0.03). There was however no main effect of Age Group (*F*(1,80)= 0.02, *p*=.889, n^2p<.01, *EMM*; Older=0.40, *SE*=0.03; Younger=0.40, *SE*=0.03), and no interaction between these factors (*F*(1,80)=1.27, *p*=.263, n^2p=.02). A Bayes factor was calculated for an independent t-test that compared the difference in *Br* between Reward conditions (Reward *Br* minus No Reward *Br*) across the two Age Groups. The Bayes factor provided anecdotal evidence in support for the null hypothesis that response bias was equal across the Age Groups ($BF_{10} = 1/2.501$).

For RTs, a mixed ANOVA (Table 2.4) with the within-subjects factors Reward (Reward vs. No Reward test) and Word Type (Old vs. New word) and the between subjects factor Age Group (Young vs. Old) showed no main effects of Reward condition (*EMM*; Reward=1208, *SE*=33, No Reward=1202, *SE*= 33) nor Word Type (*EMM*; Old word=1183, *SE*= 32; New word=1226, *SE*=35) on Response Times. There was however a significant main effect of Age Group (*EMM*; Older=1380, *SE*=44; Younger=1029, *SE*=44), with Older adults being slower overall. There was no three-way interaction, nor were there any two-way interactions.

So, to summarise, the Reward manipulation enhanced performance on the first tests only in the Younger adults, who showed improved discrimination between Old and New Words when rewarded compared to when they were not rewarded.

2.4.2 Surprise Foil Test phase

On the final foil test, Older and Younger adults did not significantly differ in their ability to correctly reject New Words (*t*(80)=1.27, *p*=.21, *d*=0.28; Younger=0.76, *SD*= 0.17; Older= 0.80 *SD*=0.14; BF_{10} = ½.17), nor did they differ in general response bias (*Br* = false alarm / (1-summary *Pr)*; *t*(80)= 1.21, *p*=.229, *d*= 0.27; mean *Br*: Older= 0.38, *SE*=0.04; Younger=0.49, *SE*=0.02; BF_{10} = ½.30).

2.4.2.1 LOP manipulation effects on subsequent foil recognition

First, I investigated whether recognition of foils on the final test differed depending on whether they had been shown intermixed with deeply encoded versus shallowly encoded old words on the first recognition test, and whether this effect was reduced in older adults, as had been previously found (Jacoby, Shimizu, Velanova, et al., 2005). Discrimination performance for the two Age Groups and LOP conditions is shown in Figure 2.3.

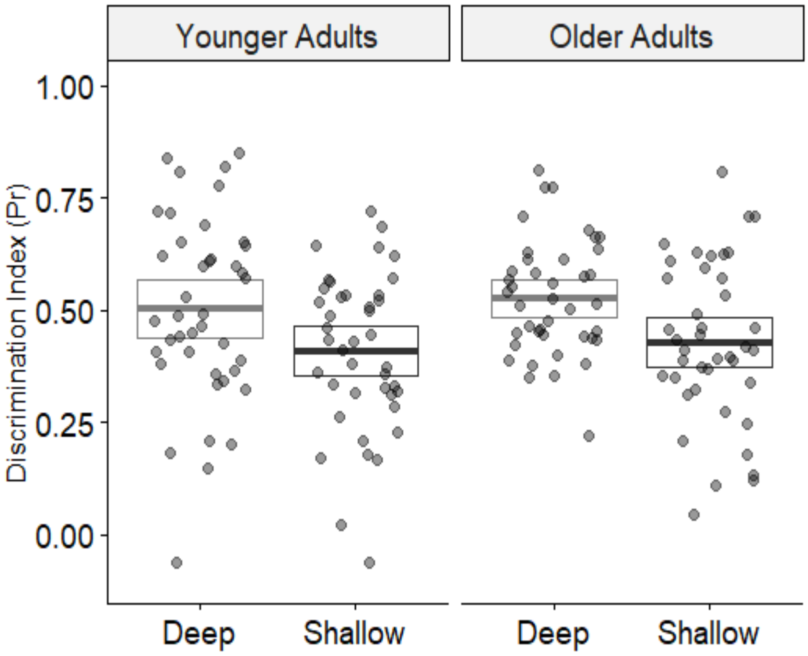


Figure 2.3. Discrimination scores (*Pr*) for Deep and Shallow foil word recognition on the final test, split by Age Group. The dot plot shows each individual score, with the mean value and confidence interval plotted as a centre line and box respectively. Scores within groups have been randomly scattered across the x-axis for visualisation purposes.

A mixed ANOVA with the within subjects factors LOP (Deep vs. Shallow) and the between subjects factor Age Group (Young vs. Old) showed a main effect of LOP on foil *Pr* ($F(1,80)=35.87, p<.001, n^2p=.31$, *EMM*; Deep = 0.51, $SE=0.02$, Shallow = 0.42, $SE= .02$) with better discrimination for Deep compared to Shallow foils. There was no main effect of Age Group ($F(1,80)= 0.40$, $p=.531$, $n^2p<.01$, *EMM*; Older = 0.48, $SE=0.03$; Younger =0.46, $SE=0.03$), nor an interaction between LOP and Age Group ($F(1,80)=0.02$, $p=.892$, $n^2p<.01$). To confirm that the effect of LOP on discrimination performance was indeed equal across Age Groups, a Bayes factor was calculated for an independent t-test that compared the difference in *Pr* between LOP conditions (Deep foil *Pr* minus Shallow foil *Pr*) across the two Age Groups. The Bayes factor provided substantial evidence in support for the null hypothesis of no difference in the effect of LOP on discrimination between the Age Groups (BF_{10} = 1/4.315). Thus, paralleling the first test, both Age Groups were equally affected by the LOP manipulation in terms of their ability to discriminate between old and new items.

Participants showed a more conservative response bias (i.e. a reduced tendency to guess "old" when uncertain) in the Shallow compared to the Deep test (main effect of LOP: *F*(1,80)=46.90*, p*<.001*,* n^2p=.37, *EMM*; Deep= 0.50, *SE=0.*03, Shallow= 0.30, *SE*= 0.02). There was no main effect of Age Group (*F*(1,80=<.001, *p*=.994, n^2p<.01, *EMM*; Older = 0.40, *SE=0.*03; Younger =0.40, *SE=0.*03) nor interaction with Age Group (*F*(1,80)=0.43*, p*=.516, n^2p<.01). To confirm that the LOP effect on response bias was equal across Age Groups, a Bayes factor was calculated for an independent t-test that compared the difference in *Br* between LOP conditions (Deep *Br* minus Shallow *Br*) across the two Age Groups. The Bayes factor provided substantial evidence in support for the null hypothesis of no difference in LOP effect on response bias between the Age Groups (BF_{10} = 1/3.61).

For RTs on the surprise foil test, Older adults responded significantly slower (M=1311ms, SE= 50) than Younger adults (M=999ms, SE= 37) to New Words (t(80)=4.96, p<.001, d= 1.10; BF_{10} = 4042.27). Older adults were also generally slower in recognising foil words from both the Deep (1314, *SD*=345) and the Shallow (1320, *SD*=435) tests, compared to Younger adults (Deep=912, *SD*=158; Shallow=925, *SD*=187). This pattern was confirmed by a mixed ANOVA with the within subjects factor LOP (Foils previously shown in the Deep vs. Shallow test) and the between subjects factor Age Group (Young vs. Old), which showed only a significant main effect of Age Group (F(1,80)= 41.09, *p*<.001, n^2p=.34, *EMM*; Older=1317, *SE*=44, Younger =918, *SE*=44), and no main effect of LOP condition (F(1,80)= 0.15, p=.703, n^2p<.01, *EMM*; Deep =1113, *SE*=30, Shallow=1122, *SE*=37) and no interaction between LOP and Age Group (*F*(1,80)=0.03, *p*=.874*,* n^2p<.01).

2.4.2.2 Reward manipulation effects on subsequent foil recognition

The next analysis was conducted to investigate whether recognition of foils on the final test differed depending on whether they had previously been shown in a recognition test where performance was rewarded or not rewarded, and whether this potential reward effect was affected by ageing. Discrimination performance for the two Age Groups and Reward conditions is shown in Figure 2.4.

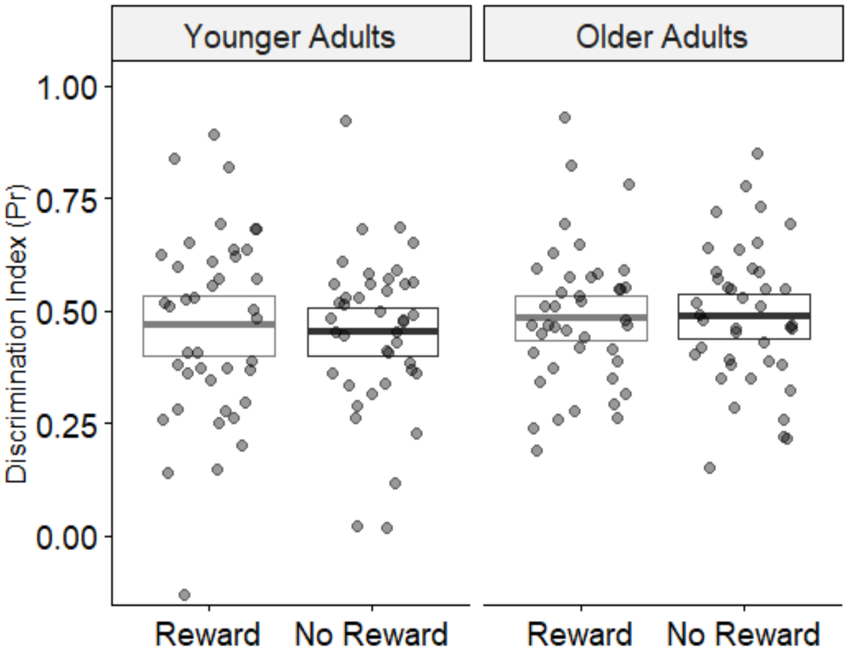


Figure 2.4. Discrimination scores (*Pr*) for the Reward and No-Reward foil word recognition on the final test, split by Age Group. The dot plot shows each individual score, with the mean value and confidence interval plotted as a centre line and box respectively. Scores within groups have been randomly scattered across the x-axis for visualisation purposes

A Reward condition x Age Group ANOVA showed no main effect of Reward condition on *Pr* (*F*(1,80)=0.08*, p*=.778*,* n^2p<.01, *EMM*; Reward=0.48, *SE*=0.02, No Reward=0.47, *SE*=0.02) and no effect of Age Group (*F*(1,80)= 0.54, *p*=.466, n^2p<.01, *EMM*; Older=0.49, *SE*=0.03; Younger=0.46, *SE*=0.03), nor an interaction between these factors (*F*(1,80)=0.24,

p=.624, n^2p<.01). To provide complementary evidence that discrimination performance was indeed similar across Age Groups, a Bayes factor was calculated for an independent t-test that compared the difference in *Pr* between Reward conditions (Reward *Pr* minus No Reward *Pr*) across the two Age Groups. The Bayes factor provided substantial evidence in support for the null hypothesis that the effect of Reward on discrimination was not different between the Age Groups (BF_{10}=1/ 3.914).

Older adults were also generally slower in recognising foil words in the surprise test that were previously shown in both the Reward (*M*=1337, *SD*=345) and the No Reward (*M*=1315, *SD*=383) tests, compared to Younger adults (Reward: *M*=914, *SD*=173; No Reward: *M*=898, *SD*=153). This pattern was confirmed by a mixed ANOVA with the within-subjects factor Reward condition (Foils previously shown in the Reward vs. No Reward test) and the between subjects factor Age Group (Young vs. Old), which showed only a significant main effect of Age Group (F(1,80)= 53.08, p<.001, n^2p=.399, *EMM;* Older = 1326, *SE*= 41, Younger= 906, *SE*= 41). There was no main effect of Reward condition (F(1,80)= 0.66, p=.419, n^2p<.01, *EMM*; Reward= 1126, *SE*= 30, No Reward = 1106, *SE*= 32) and no interaction between Reward and Age Group (*F*(1,80)=0.02, p=.899, n^2p<.01).

Thus, analysis of reaction times on the foil test only revealed a general slowing of responses in the Older compared to Younger group, but no foil condition differences nor interactions between foil conditions and Age group.

2.4.3 Correlations between older adults' demographics and the LOP foil effect

In a final analysis of the Surprise test data, I investigated whether the source-constrained retrieval effect (i.e. the effect of LOP on foil recognition) was related to age or level of education within the older sample. Since I had not found an age-related impairment in source-constrained retrieval as described in prior research (Jacoby, Shimizu, Velanova, et al., 2005), I wanted to assess whether this might be related to demographic variables of relevance (for example, perhaps only the oldest or less educated adults in our sample showed a reduced source-constrained retrieval effect). These issues were examined by calculating the difference in foil discrimination (*Pr*) on the final test between the Deep and Shallow conditions, and correlating this difference with age and numbers of years in education within the Older adults only. There was no significant relationships between age (r_s =.045, p=.783; $BF_{10, (r \text{ width}=1)}$ = 1/4.76) or years in education (r_s =.178, p=.271; $B_{10, (r \text{ width}=1)}$ = 1/4.98, N=40) for both due to one participant missing demographic data) and the LOP source-constrained retrieval effect, and the Bayes Factor indicated support for the null hypothesis of no association between variables in both cases. However, these results should be interpreted cautiously since our study design and sample size was not optimised for detecting such correlations.

2.4.4 Motivation, effort and strategy use as a function of Reward

Since rewards only enhanced performance for the Young adults but not the Older Adults on the first test, a supplementary analysis tested if there were differences in Younger and Older adults' reported motivation and effort for the Rewarded and Non-Rewarded conditions, as indicated by participants self-reports in the post-experimental questionnaire (see Table 2.5 for abbreviated questions asked).

Table 2.5. *Experiment 1 post-experimental questionnaire.*

Question	Post Experiment Questionnaire	Response Options
1	Did you anticipate the final Surprise Test?	Y/N
2	Did you use any techniques to help you learn the words in the rewarding learning task? If so briefly tell us about these techniques.	Y/N
3	Did you use any techniques to help you learn the words in the standard (no reward) learning task? If so briefly tell us about these techniques.	Y/N
4	Can you rate how motivated you were to win the maximum amount of money during the rewarding learning task?	scale 1 to 10
5	Can you rate how much effort you put into learning the words during the rewarding learning task?	scale 1 to 10
6	Can you rate how motivated you were to win the maximum amount of money during the rewarding memory test?	scale 1 to 10
7	Can you rate how much effort you put into recognising the words during the rewarding memory test?	scale 1 to 10
8	Can you rate how much effort you put into learning the words during the standard (no reward) learning task?	scale 1 to 10
9	Can you rate how much effort you put into recognising the words during the standard (no reward) memory test?	scale 1 to 10
10	Please write down how many years you have been in education	…

Note: Y/N means Yes/No response options, For the scale of 1 to 10: 1 represents low endorsement of the item, with 10 being the highest level of endorsement.

In the Rewarded phases, the Older adults reported being less motivated (Table 2.6) to gain the monetary reward in both the study ($t(80)=7.59, p<.001, d=1.67$; $BF_{10}=1.386^{e+08}$) and first recognition test phases ($t(76.13)=5.87, p<.001, d=1.35$; $BF_{10}=121875.99$) compared to the Younger adults. Yet, both groups reported putting in a similar amount of effort in the rewarded study ($t(70.59)=.218, p=.828, d=0.05$; $BF_{10}=1/4.26$) and test phases ($t(71.29)=.439, p=.662, d=0.01$; $BF_{10}=1/3.99$). However, in the standard non-rewarding phases, Younger adults reported lower effort to learn words ($t(80)=3.93$, $p<.001$, $d=0.87$; $BF_{10}=133.04$) and made less effort to retrieve words ($t(80)=3.15$, $p=.002$, $d= 0.70$; $BF_{10}=15.08$) than Older adults.

When examining effort within each age group, Younger adults reported using less effort to learn words in the Non-Rewarded study phase compared to the Rewarded study phase (*t*(40)=2.68, *p*=.011, *d*=0.40 , BF_{10} =3.79), whereas Older adults reported the opposite effect, putting more effort into learning the non-rewarded than rewarded words (*t*(40)=2.65, *p*=.011, *d*=0.35 , BF_{10} =3.63). Similarly, Younger adults reported making less effort to recognise words in the Non-Rewarded test compared to the Rewarded test (*t*(40)=3.12, *p*=.003, *d*= 0.40; BF_{10} =10.33), whereas there was no significant difference in effort across the tests for Older adults (*t*(40)=1.90, *p*=.065, *d*= 0.18; BF_{10} =1/1.17).

Since it has been suggested that older adults are less likely to spontaneously generate and apply strategies in memory tasks and that this may contribute to reduced memory performance (Morcom, 2016), we also tested whether there were differences in strategy use between Younger and Older adults during the Rewarded and Non-Rewarded study phases (Table 2.6). Young adults reported that they used some type of technique to learn words more often than Older adults in both the Rewarded (X^2(1)= 8.11, *phi*=.31, *p*=.004) and Non-Rewarded study phases (X^2 (1)= 4.04, *phi*=.22, *p*=.044). Participants that reported using strategies were also asked to describe their strategies open-endedly. These responses were not possible to analyse quantitatively, but there were some similar themes. The majority of strategy users in the younger adult sample described strategies involving elaboration in some form. For example, participants reported associating words to each other, to themselves or an image, with some reporting integrating the words into a story. A smaller proportion of the younger adults used non-elaborative strategies such as focussing on the words individually or rehearsing the words in working memory. Out of the older adults who reported using strategies, a similar proportion also used an elaborative memory strategy; primarily forming associations between words, themes or to themselves, again with a small proportion reporting using sub-vocal rehearsal as a strategy. Some older adults who reported using a strategy did however encounter difficulties with successfully implementing these self-generated strategies, and reported that the strategy became too difficult to use for the whole word list.

Thus, the questionnaire self-reports converged with the memory performance measures in suggesting that older adults were less affected by rewards than younger adults, and also showed that younger adults were more likely than older adults to intentionally use strategies to facilitate learning when an encoding task was not provided.

Table 2.6. *Self-report measures of motivation, effort and strategy use for the Study and Test 1 phases in the Reward and Non-Reward blocks, split by Age group.*

Reward Manipulation	Self-report measures	Younger Adults		Older Adults	
		Study Phase	Test Phase	Study Phase	Test Phase
Reward phase	Motivation to gain reward	7.02 (2.02)	7.02 (2.02)	3.66 (2.00)	4.05 (2.54)
	Amount of effort	7.46 (1.61)	7.93 (1.42)	7.56 (2.37)	8.10 (2.05)
	Strategy use	83%	-	54%	-
No Reward phase	Amount of effort	6.78 (1.80)	7.27 (1.83)	8.27 (1.63)	8.44 (1.52)
	Strategy use	68%	-	46%	-

Note: Motivation and effort values shown are Means (SD), and were collected on a scale between 1-10, with a higher score indicating more of the item measured. “Strategy use” refers to percent of participants who reported using a strategy to learn words in the study phases. See Table 2.5 for questionnaire items.

2.5 Discussion

The current study investigated whether older adults show deficits in strategic retrieval processes, as has been found in previous research (Cohn, et al. 2008; Kirchhoff, et al., 2012; Morcom, 2016). Specifically, I aimed to replicate previous findings that older adults are less likely than young adults to spontaneously re-implement processes from a prior encoding task during a subsequent recognition test in order to constrain their retrieval search towards a particular source (Jacoby, Shimizu, Velanova, et al., 2005). I also extended on previous research by exploring the effects of monetary rewards on incidental encoding processes engaged during retrieval attempts, in order to investigate if age deficits in strategic retrieval processing were accompanied by reduced reward influences on memory within the same sample.

The main findings showed no influence of rewards on incidental encoding during retrieval attempts in either age groups. However, there was evidence that both age groups engaged in strategic (LOP) source-constrained retrieval to a similar extent. Thus, I unexpectedly found that the older adults appeared to have intact ability to strategically reinstate encoding processes during retrieval, conflicting with prior findings (Jacoby, Shimizu, Velanova, et al., 2005).

2.5.1 Spontaneous LOP reinstatement

The older and younger groups were comparable on general memory performance and showed equivalent effects of the encoding LOP manipulation (Craik & Lockhart, 1972) during the initial encoding task, as assessed with a first old/new recognition test that showed that both groups benefited equally from a 'deep', semantically oriented task compared to a 'shallow', perceptually oriented tasks. Such equivalent effects of LOP encoding task manipulations on recognition accuracy across age groups have been found previously (Grady, McIntosh, Rajah, Beig, & Craik, 1999; Jacoby, Shimizu, Velanova, et al., 2005), even though older adults typically also show a general reduction in recognition discrimination ability (Fraundorf et al., 2019), which was not found in this older group. Furthermore, in the surprise foil test that followed, both groups showed equivalent enhanced encoding of foils that had been tested together with old items from the deep encoding task compared to foils that had been tested together with old items from the shallow encoding task. This pattern suggests that both age groups spontaneously and strategically altered their retrieval processing, in line with the source-constrained retrieval theory (Jacoby, Shimizu, Daniels, et al., 2005), producing differential incidental encoding of foils. In the memory-for-foils paradigm, enhanced subsequent recognition for the deep over the shallow foil words is thought to be caused by participants spontaneously reinstating deep vs. shallow encoding processes during the initial tests. That is, this strategy thus involves focusing on semantic aspects of stimuli ("deep" processing) in a recognition test if the initial encoding task involved semantic judgements, versus focusing on perceptual aspects of stimuli ("shallow" processing) if the initial encoding task involved perceptual judgements. By reinstating such different processing modes, the foils become more deeply encoded in the deep test than the shallow test. The claim that people are able to strategically direct retrieval towards different types of information at test has been supported by several different lines of evidence (Alban, 2010, 2013; Alban & Kelley, 2012; Halamish et al., 2012). For example, EEG and fMRI studies have shown with neural evidence that reinstating a "deep" processing mode (indicated by left PFC activity and alpha oscillations) during recognition attempts is what drives the differences in foil encoding, and other behavioural evidence also supports this view (see Danckert, et al., 2011; Gray & Gallo, 2015; Jacoby, Shimizu, Daniels, et al., 2005; Vogelsang et al., 2016, 2018).Although there are other potential explanations for the source-constrained retrieval effect, such as for example foils being more distinctive in a "deep" than "shallow" test, such alternatively explanations have not been widely supported in prior literature.

Prior research has suggested that intact recognition test performance in older adults may be due to a reliance on automatic familiarity over more controlled recollection of context (Koen & Yonelinas, 2016; Yonelinas, 2002). However, if the older adults had relied on automatic memory processes during the first recognition test, I would not have expected to see any difference between foils in terms of subsequent memory in this group. Instead, it appeared that they were just as likely as the young group to engage strategic retrieval processes to facilitate recollection of the study context in accordance with the source-constrained retrieval theoretical framework (Jacoby, Shimizu, Velanova, et al., 2005). These results therefore conflict with prior findings from a very similar experimental design that also compared young university students against older members of the community and found no differences between deep and shallow foils in the older participants (Jacoby, Shimizu, Velanova, et al., 2005), and are also inconsistent with other findings of age-related impairments to strategic retrieval processing (reviewed in Morcom, 2016). It is unclear exactly what caused this inconsistency. The older sample was of similar age (mean age 73, range 64-85) to those in prior relevant studies (e.g. Jacoby, Shimizu, Velanova, et al., 2005, where the mean age was 75 and range 61-87). Furthermore, there was no indication of a reduction in the source-constrained retrieval effect (i.e. the difference between 'deep' and 'shallow' foil recognition discrimination) as a function of age within the older sample, despite the relatively large range between our youngest 'old' and oldest 'old' participant.

However, it is well known that ageing-related memory problems can be highly variable across individuals of the same age, and there are many people who have intact memory function even at a very old age (Nyberg et al., 2012). One possibility is that this older sample was unusually high performing because of demographic reasons – they were mostly highly educated, and all were socially active members of the University of the Third Age, which is an organisation for lifelong learning that encourages retired or semi-retired people to share their skills and knowledge with each other. High levels of education is typically associated with intact cognitive performance in old age, a correlation between years in formal education and the SCR effect was not observed in this sample (although there was a restricted range of values for education, so this lack of association should be interpreted cautiously). Education is not the only important factor for healthy cognitive ageing however, since being socially active and engaging in mentally challenging activities have also been found to independently predict high cognitive functioning in old age (Chan et al., 2018; Nyberg & Pudas, 2019). Members of social/educational groups such as the U3A may be particularly likely to have a lifestyle that promotes brain health and intact cognitive function in old age. Although I did not intentionally aim to sample older adults with particularly high memory functioning, I may nevertheless have done so by recruiting only from the U3A. Consistent with this account, these participants did not show a general deficit in memory performance as found in the older sample in Jacoby, Shimizu, Velanova, et al. (2005), where older participants' recognition accuracy was reduced overall across initial and final recognition tests.

It should also be noted that reductions in strategic retrieval processing in older age have not always been found in prior research. As reviewed in Morcom (2016), older adults sometimes engage strategic retrieval control when required to do so by the task (Duverne, Motamedinia, & Rugg, 2009). This finding suggests that ageing may not lead to permanent impairments in strategic retrieval processes, but rather that older people tend not to engage such processes spontaneously in item recognition tasks (see also Morcom & Rugg, 2004). It is possible that subtle aspects of this procedure could have encouraged older participants to use source-constrained retrieval strategies more effectively than in prior studies. For example, it is known that experimenter and participant expectations can have strong effects on cognitive test results, such as in the case of stereotype threat where older adults perform worse when it is implied that their memory is likely to be impaired (Krendl et al., 2015; Lamont, Swift, & Abrams, 2015). Such expectancy effects may be more likely when a younger experimenter tests older adults (McDaniel et al., 2008). In this study, U3A members who were of similar age as the older group collected data from that group, which may have reduced expectancy effects. Great care was taken to ensure that the computerised task requirements and instructions would be suitable for the older sample, which may have enabled this group to perform well. Thus, the equivalent performance of the two age groups may be due to individual cognitive and/or situational factors.

2.5.2 External reward effect on SCR

Moreover, in this study, I added a novel manipulation to investigate whether external rewards would modulate incidental encoding of foils during a recognition test, and whether such effects would differ between young and older adults, as might be expected based on prior findings of reduced reward-sensitivity in older age (e.g. Geddes, et al., 2018; but see Mather & Schoeke, 2011). However, the monetary reward manipulation only affected recognition of words that had been intentionally encoded, and only in the young group. The younger adults showed enhanced recognition on the first test for old words that had been intentionally encoded in a study phase with reward compared to those encoded in a standard, non-rewarded study phase, whereas the older group showed no difference between reward and non-rewarded conditions on the first test. Consistent with these performance differences, younger adults reported higher levels of motivation for the rewarded study and test phases, coupled with lower effort to learn and retrieve non-rewarded words compared to older adults. Thus when compared to older adults, younger adults did not only seem to increase their encoding efforts because of a strong motivation to win a monetary reward, but they also reduced their retrieval efforts when there was no reward, perhaps to minimise interference of non-rewarding words (Hennessee et al., 2017), or simply due to finding the tasks intrinsically less motivating than the older group (Murayama & Kuhbandner, 2011). Older adults however showed high recognition performance for both the rewarded and non-rewarded conditions, indicating that their memory ability was equal to the younger adults, but they were not incentivised by the monetary reward. Instead they reported being less motivated by monetary reward than the young adults, in line with the literature on the 'undermining effect' of monetary rewards (for review see Deci, Koestner, & Richard, 2001). The older adults were instead perhaps more motivated by an internal drive to succeed and perform well, as indicated by their high effort levels in the non-rewarded phase.

In contrast to the reward effects on the first test, neither group showed a difference in recognition accuracy on the final surprise test for foils that had been shown in a rewarded versus non-rewarded recognition test. I expected to find enhanced incidental encoding of rewarded compared to non-rewarded foils, since previous research has shown that incidental encoding is enhanced for stimuli associated with reward (Adcock et al., 2006; Mather & Schoeke, 2011; Spaniol, Bowen, Wegier & Grady, 2015). More generally, prior studies have found that factors that enhance intentional encoding also enhance incidental encoding of foils in a recognition test, for example self-referential processing, imagery, full attention, and survival processing (Bergström et al., 2015; Danckert, et al., 2011; Dudukovic et al., 2009; Nairne et al., 2015). I hypothesised that rewards during a recognition test could enhance encoding of foils either directly through automatic, dopamine-mediated brain mechanisms (eg. Bäckman et al., 2010) or more indirectly by increasing participants' use of strategic retrieval processes (eg. Halsband et al., 2012), which could have resulted in more attentive/deeper processing of foils and thereby enhanced encoding. However, the results showed no support for either of these accounts. Since the effect of the reward manipulation on intentional encoding was rather subtle, the manipulation may not have been strong enough to produce differential encoding of foils. For example, even if participants sometimes used "deep" strategies to enhance encoding of words in the rewarded compared to the non-rewarded study phase, they may have used such strategies inconsistently (e.g. switching strategies across items within a study phase, or using similar strategies in both rewarded and non-rewarded study phases). Therefore, the reward study manipulation may not have provided a clear source context that participants could use to constrain retrieval towards during the subsequent test, thus leading to similar retrieval processing in both the rewarded and non-rewarded test, and no difference in incidental foil encoding. Based on the current null results of rewards on foil encoding in both groups, firm conclusions regarding the relationship between reward-induced motivation and strategic control processing as modulators of memory, and how these factors and their relationship are influenced by ageing cannot be drawn.

2.6 Conclusions

In sum, the main finding in this experiment was fully intact source-constrained retrieval processing in a group of socially active, mostly highly educated older adults who are members of a lifelong learning organisation. The older group had equal memory accuracy to a group of university students who were on average 50 years younger, and showed equivalent engagement of complex retrieval strategies. Although I cannot provide a conclusive explanation for why the older sample of adults had intact memory performance when so many prior studies have found ageing-related impairments on similar memory tasks, I believe this is an important demonstration that ageing-related memory decline is far from inevitable and universal. In contrast to the first test, neither group showed a difference in foil memory on the final surprise test for foils that had been shown in a rewarded versus non-rewarded recognition test. I expected to find enhanced encoding of rewarded compared to non-rewarded foils, since previous research, has shown that incidental encoding is enhanced for stimuli associated with reward (Adcock et al., 2006; Mather & Schoeke, 2011; Spaniol et al., 2015). Furthermore, many prior studies have found that factors that enhance intentional encoding also enhance incidental encoding of foils in a recognition test, for example self-referential processing, imagery, full attention, and survival processing (Bergström et al., 2015; Danckert et al., 2011; Dudukovic et al., 2009; Nairne et al., 2015). However, since the effect of the reward manipulation on intentional encoding was rather subtle, it may not have been strong enough to produce differential encoding of foils.

2.7 Summary

In this chapter I expanded on the prior literature on strategic retrieval processes in older adults and contributed a novel exploration of how monetary rewards affect both intentional and incidental encoding in young and older adults. Through the citizen science collaboration, I was able to reduce performance biases caused by experimenters' and participants' expectations about ageing and memory, to investigate the effects of healthy ageing on strategic as well as reward-based source constrained retrieval effects. Here, I demonstrate the versatility of the memory-for-foils paradigm by using it both to contribute directly to new empirical findings as well as adapting this paradigm to investigate reward effects on memory. My findings show novel evidence for intact spontaneous strategic reinstatement in healthy older adults as well as differences in reward processing between intentional and incidental encoding. In the next chapter, I further modified the memory-for-foils paradigm to combine it with the value-directed remembering paradigm (Castel, 2007), in order to examine the effects of rewards on recognition more.

Chapter 3: Value-directed recognition; a points-based manipulation.

3.1 Overview

This chapter built on the findings from Chapter 2, where I found that older adults, contrary to previous suggestions (Jacoby, Shimizu, Velanova, et al., 2005), were able to strategically use source-constrained retrieval to aid recognition, resulting in equivalent incidental encoding of foils as younger adults. Additonlly, only the younger adults enhanced their recognition performance after intentional encoding with rewards, whereas the older adults reported not being motivated by a monetary reward and did not show any effects of reward on memory. Neither group were influenced by rewards enough to alter incidental foil encoding rates, and in Chapter 2 I suggested some possible reasons why the external reward did not alter incidental encoding rates. In this chapter, I aimed to address these issues.

In Experiment 2, I modified the memory-for-foils paradigm further to incorporate elements from a second paradigm, value-directed remembering (Castel, 2007; Castel, Benjamin, Craik, & Watkins, 2002) where assigning arbitrary values to different stimuli during encoding leads to encoding prioritisation. To my knowledge, such manipulations of subjective value have not previously been used to alter retrieval processing during a memory test. In this chapter, I explored if value presented during recognition attempts can alter incidental encoding of novel foil items. This chapter thus bridges the literature on value-directed remembering with literature on memory-for-foils.

3.2 Experiment 2

An important aspect of human memory is the ability to selectively attend to and adapt in response to important and highly salient information. Highly salient information attracts attention and is typically more valuable to remember (Anderson, 2016; Castel et al., 2013). Rewards increase the likelihood of successful remembering (Shohamy & Adcock, 2010), both through more short term strategic intentional (Castel et al., 2007, 2013; Cohen et al., 2014) and incidental learning (Adcock et al., 2006; Hauser et al., 2019; Mather & Schoeke, 2011; Yan, Li, Zhang, & Cui, 2018). Items judged in associated to; or in anticipation of high rewards are incidentally learnt and subsequently recognised more. This has demonstrated the powerful effect of rewards on cognitive processes such as attention (Anderson, 2016; Chiew & Braver, 2016; Miendlarzewska et al., 2016) and task monitoring and cognitive control (Chiew & Braver, 2016). Expected rewards have been seen to lead to intentional, strategic prioritisation of highly valuable information (Castel et al., 2002, 2007, 2013; Cohen et al., 2014; Hennessee et al., 2017). For example, through disproportionately assigning more study time to high value items (Ariel et al., 2015; Castel et al., 2013; Robison & Unsworth, 2017), or alternatively through selectively re-studying high value items (Castel et al., 2013, Middlebrooks & Castel, 2018).

The value-directed remembering paradigm (Castel, 2007; Castel et al., 2002) has frequently been used to investigate the strategic prioritisation of resources to maximise reward earnings. In this paradigm, during encoding tasks, study items (e.g. words) are assigned to a point value so that participants are promised either high, low or even minus points if they later remember those items in a subsequent test (Castel, 2007; Castel et al., 2002). It has been found repeatedly that high value items are intentionally encoded and remembered more than low or negative value items, and modifications of this basic task have been used to examine different aspects of spontaneous strategy use during encoding (Cohen et al., 2014; Hennessee et al., 2017; Hennessee, Patterson, Castel, & Knowlton, 2019; Robison & Unsworth, 2017, Middlebrooks & Castel, 2018) with particular interest in prioritisation and strategy use in ageing (Ariel et al., 2015; Castel et al., 2007, 2011, 2013).

Value-directed remembering studies have primarily used free recall to assess encoding success. To my knowledge, only one prior study has examined value-directed remembering with a recognition test (DeLozier & Rhodes, 2015), and there are limited studies manipulating value during recognition retrieval attempts (Han, Huettel, Raposo, Adcock, & Dobbins, 2010; Yan et al., 2018). Building on this, no studies to my knowledge have manipulated value during recognition attempts to try to alter retrieval processing with the view that this will influence incidental foil encoding. There has been evidence however showing that retrieval processes can be altered in anticipation of differing levels of rewards (Halsband et al., 2012), but the effect on subsequent retrieval due to incidental encoding has not been shown yet. Along with altering retrieval processes and allocation of cognitive resources, the episodic memory system can adaptively assign value and importance to information retroactively (Patil, Murty, Dunsmoor, Phelps, & Davachi, 2017). The ability to retroactively assign value to information is important, coupled with as discussed earlier, the ability for rewards to incidentally increase learning through capturing attention (Anderson, 2016) and increasing cognitive monitoring (Chiew & Braver, 2016) provides a potentially powerful memory tool suggesting that altering incidental encoding is reasonably justified..

This study therefore aimed to extend on previous research by combining elements from the literature on value-directed remembering (Castel et al., 2007) with elements from the retrieval processing literature (Buckner et al., 2001; Danckert et al., 2011; Halsband et al., 2012; Jacoby, Shimizu, Daniels, et al., 2005) and the incidental reward based learning literature (Adcock et al., 2006; Hauser et al., 2019; Mather & Schoeke, 2011). This experiment followed the three memory-for-foils phases, with an initial encoding phase, followed by a recognition task and a final surprise recognition task which measures the rate of incidental encoding occurring during the initial test phase. Unlike in most memory-for-foils tasks, the initial encoding phase was not of interest here and only used to make sure the engagement in the first recognition task was sufficient. The first recognition task was used to examine the effect of the presence of reward on recognition processes, similarly to the use of rewards to alter other judgment processes to increase incidental encoding (Adcock et al., 2006; Mather & Schoeke, 2011). Therefore, participants initially took part in a shallow judgment task (the encoding phase). Then when participants were asked to recognise and correctly identify these items in the recognition task, they were told some items would hold a high value and other a low value. The point values would be summed to provide a total score for each participant, with points being allocated to both old and new items. Participants were told to try to achieve the highest total point score to win prizes. The best technique, though not instructed to participants, would be to ignore processing the low value items during the recognition task as there was no negative scoring and prioritise the high value items. Finally, after the 'value-directed' recognition phase, there was a surprise recognition phase. During this test there was no value linked to the foil or new words. This phase enabled the rate of incidental encoding to be examined, to see if high compared to low value words during the initial recognition test were processed differently, with the aim to maximise point winnings.

Therefore, overall Experiment 2 aimed to extend on the current literature on retrieval processing, by investigating both strategic prioritisation of resources to high value items as well as more automatic influences of high value on incidental encoding rates. Rewards have been seen to alter episodic memory both through intentional and incidental processes, through automatic and strategic mechanisms as reviewed in Chapter 1. Therefore, I expected that participants would prioritise processing of high value items during the recognition task, potentially enhancing incidental encoding of high value novel foil words. Recognition performance on the first test was examined to investigate the effect of point values on initial retrieval processing, as measured by reaction times, recognition accuracy/discrimination performance, and response bias.

I then examined how point values introduced during retrieval affected incidental encoding, as measured by subsequent foil recognition accuracy and reaction times. However, as this was a novel manipulation of value-directed retrieval, it was unclear whether previously described effects of high value points on prioritisation during intentional encoding would also be found for retrieval and incidental encoding rates. Tentatively, I predicted that if participants strategically prioritised high value items during the initial test, this may increase the accuracy of recognition decisions as a result of more effortful and careful processing of high compared to low value items. However, if participants were unable to strategically enhance recognition accuracy, prioritisation may instead be reflected in changes to reaction times and response biases, without any changes in recognition accuracy. If participants' prioritisation of high value items (seen at test 1) enhanced incidental foil encoding, this would lead to more accurate and potentially faster recognition of high compared to low value foils on the final surprise test.

3.3 Methods

3.3.1 Participants

Results are based on 86 native English-speaking student participants (17 Male; *M* age=18.97 years, *SD*=1.2, range = 18-25). This sample size enabled me to achieve >0.9 power to detect an effect size of Cohen's *d*=0.5 at two-tailed α=.05 (or alternatively ~0.8 power to detect an effect size of Cohen's *d*=0.3 at two-tailed α=.05). The participants were recruited from the University of Kent Psychology participation scheme where students received course credits for their time. Participants had normal or corrected to normal vision (colour-blindness was not tested and only self-reported) and reported not having dyslexia and not taking any psychoactive medication. Participants could have been left or right handed. Participants had not taken part in any other experiments within this thesis.

All participants provided full written consent and were entered into a score board based on amount of points collected as part of the value manipulation during the memory test. The top scorer won £20 in Amazon voucher, and the two runners up both won £5 in Amazon vouchers. Participants were not aware of the chance to win Amazon vouchers before they signed up for the study and the procedures were approved by the University of Kent, School of Psychology Ethics Board. Participants were tested either by *L.S.* or by three final year undergraduate students who were trained as experimenters by *LS*. All testing was conducted in a face-to-face manner.

3.3.2 Materials

Similarly to Experiment 1, the experiment was programmed and run in PsychoPy (Peirce, 2007). A subset of words (84%) from Experiment 1 (Chapter 2) were re-used giving a total of 240 English words. These words were originally from the MRC psycholinguistic database (3 - 11 characters, Kucera-Francis frequency range 0-442, Wilson, 1988) and initially selected from a larger set used by Vogelsang et al., (2016, 2018). The words were randomly sorted into three lists with 80 words in each, which were counterbalanced across participants yielding 3 counterbalancing rotations.

A set of 80 emotionally neutral images from the IAPS database (Lang et al., 2005) were used in a visual search filler task (50% of the IAPS images previously used in Experiment 1, Chapter 2). Apart from a lower number of stimuli, the filler task was the same as in Experiment 1 (see 2.3.3.2), and this task was conducted once after encoding and then once after the retrieval task.

3.3.3 Design and Procedure

The procedure was modified from a memory-for-foils paradigm (Jacoby, Shimizu, Daniels, et al., 2005; Vogelsang et al., 2016) to manipulate retrieval processing based on value magnitude, inspired by value-directed remembering paradigms (Castel et al., 2007). Similar to Experiment 1, all participants completed an initial judgment task during word encoding. In contrast with Experiment 1, all participants then completed a value cued recognition task rather than a standard recognition task, with the intention to manipulate retrieval orientation during retrieval attempts. Finally, participants went on to complete a surprise recognition task for the foils from the first test, which had a similar format to in Experiment 1. The schematic in Figure 3.1 presents the three task stages. In between each stage a visual search filler task was used which followed the same protocol as in Experiment 1 (see Chapter 2.3.3.2; Figure 2.3 for filler task schematic).

As Experiment 2 bridged across different literatures, this novel paradigm design needed to be adapted from two different paradigms. The memory-for-foils paradigm commonly uses a blocked manipulation of retrieval processes across two different recognition tests (Jacoby, Shimizu, Daniels, et al., 2005), whereas value-directed remembering typically assigns points to items at encoding in an intermixed fashion (Castel et al., 2002, 2007). Importantly for this paradigm modification, on the whole during intermixed cuing, retrieval processing is adapted based on retrieval cues, with some behavioural costs to accommodate for processing switching (Johnson & Rugg, 2006; Marsh et al., 2009). To achieve a reinstated encoding strategy or create a stronger retrieval orientation towards a specific cue or context, then a blocked design would be more efficient (Alban & Kelley, 2012; Wilding & Nobre, 2001). Yet, since the main aim in Experiment 2 was to examine the effects of value on retrieval processing rather than intentional encoding or reinstated processing, I decided to use an intermixed design to increase the contrast in value between items shown within the same test, which might have been reduced in a blocked design (as used in Experiment 1, where high vs. low rewards were manipulated across two different tests). Intermixed manipulation of value is important, because points are not inherently valuable in themselves, but the contrast between a high and low value point increases the saliency of high value points (Madan & Spetch, 2012).

Another paradigm design element I had to address was the initial encoding task, this needed to be engaging but not too effective, otherwise performance on the initial recognition test may have been at ceiling. Task difficulty is important to encourage participants to try to strategically prioritise information processing, since easy tasks do not require prioritisation of resources. Prior research has shown that memory tasks need to be at least moderately difficult so that rewards and value are used to aid prioritisation of items to be remembered through task demands (Pyc & Rawson, 2009; Shigemune, Tsukiura, Nouchi, Kambara, & Kawashima, 2017) or task interest (Murayama & Kuhbandner, 2011). I therefore decided to use a shallow processing task (Craik & Lockhart, 1972) during encoding to maintain a relatively high difficulty on the subsequent recognition test, with the intention to encourage participants to enhance prioritisation of high value items.

3.3.3.1 Main Encoding and Recognition tasks

Firstly, participants were given a shallow judgment task, in which they were presented with a list of 80 words and were asked to judge whether each word contained an ‘O’ or ‘U’. Words were presented for 2000ms, and participants pressed ‘Z’ or ‘M’ keys on the keyboard within this time to indicate ‘Yes’ or ‘No’ respectively. Counterbalancing of button press handedness was not done here as RT differences between old vs. new items was not a main hypothesis of interest and the interaction with reward could still be examined. Responses were only accepted 100ms after word onset to promote engagement with the task. Each word was preceded by a 700ms fixation cross. The fixation cross duration was increased by 100ms compared to in Experiment 1, after pilot feedback that participants felt they were being rushed on this particular task given that word lists were longer.

Participants then took part in a recognition task. Participants were told that they would receive points for correctly identifying words they previously recognised from the judgement task and for correctly identifying new words, and were asked to try to score the maximum amount of points they could. Before each word, a cue indicated how many points they would receive for correctly identifying the upcoming word as either old or new. These cues indicated either Low Value ('1' point) or of High Value ('12' points), and these point values were equally split for Old and New words so that overall, 80 old words (40 High Value and 40 Low Value Old Words) were randomly intermixed with new foil words (40 High Value and 40 Low Value Foil Words). The point cue was displayed in red font within a golden coin for 90ms before the word presentation. Then, the word was presented below the coin for 2100ms and participants could respond only after 100ms of word onset to reduce the risk of random button presses. Participants pressed 'Z' or 'M' keys on the keyboard within this time to indicate 'Yes' or 'No' respectively. Before each recognition trial, a pre-stimuli fixation cross was presented for 700ms. Participants were told that their total point score would be entered into a point league board and that the top scorer would win £20.00 and the two runners up winning £5.00 each in Amazon vouchers.

Participants then received a surprise recognition test which included all of the 80 foil words (40 Low Value and 40 High Value) randomly intermixed with new words (80 New). Participants were instructed that old words could be presented in this test from either the judgment task or the recognition task, and their task was to recognise any words from the previous tasks in the experiment. Again, participants could only respond after 100ms later than the word onset. Words were shown on the screen until a response was made. Participants pressed 'Z' or 'M' keys on the keyboard to indicate 'Yes' or 'No' respectively. A pre-stimulus fixation cross was presented for 500ms before the word presentation (as the task was self-paced it was decided that a longer fixation cross was not needed here).

3.3.3.2 Visual Search filler task

A visual search task was used as a filler to prevent rehearsal between each experimental task (two times in total). Participants were presented with an image and were asked to search for a white triangle that had been superimposed onto a subset of images. As in previous tasks, participants again used the ‘Z’ or ‘M’ keys to respond ‘Yes’ or ‘No’, this time to the question, ‘Is there a white triangle on the image?’. This task lasted 1 minute in total and was the same as in experiment 1 (see Chapter 1; section 2.3.3.2 for more details of this task).

3.3.3.3 Post-Experiment tasks and Debrief

After participants had completed the experimental tasks, they were told their score and asked to complete a post experiment questionnaire about how motivated they were to win the maximum amount of points, the level of effort put into recognising the words and how difficult they found the first test. I also asked them about any strategies used during the first recognition task which they thought might help them identify words more successfully. Participants were also asked if they anticipated the surprise test phase (see Table 3.2 for abbreviated questions asked). All participants were given verbal and written debrief and were told they would be contacted via email if they won one of the top 3 prizes.

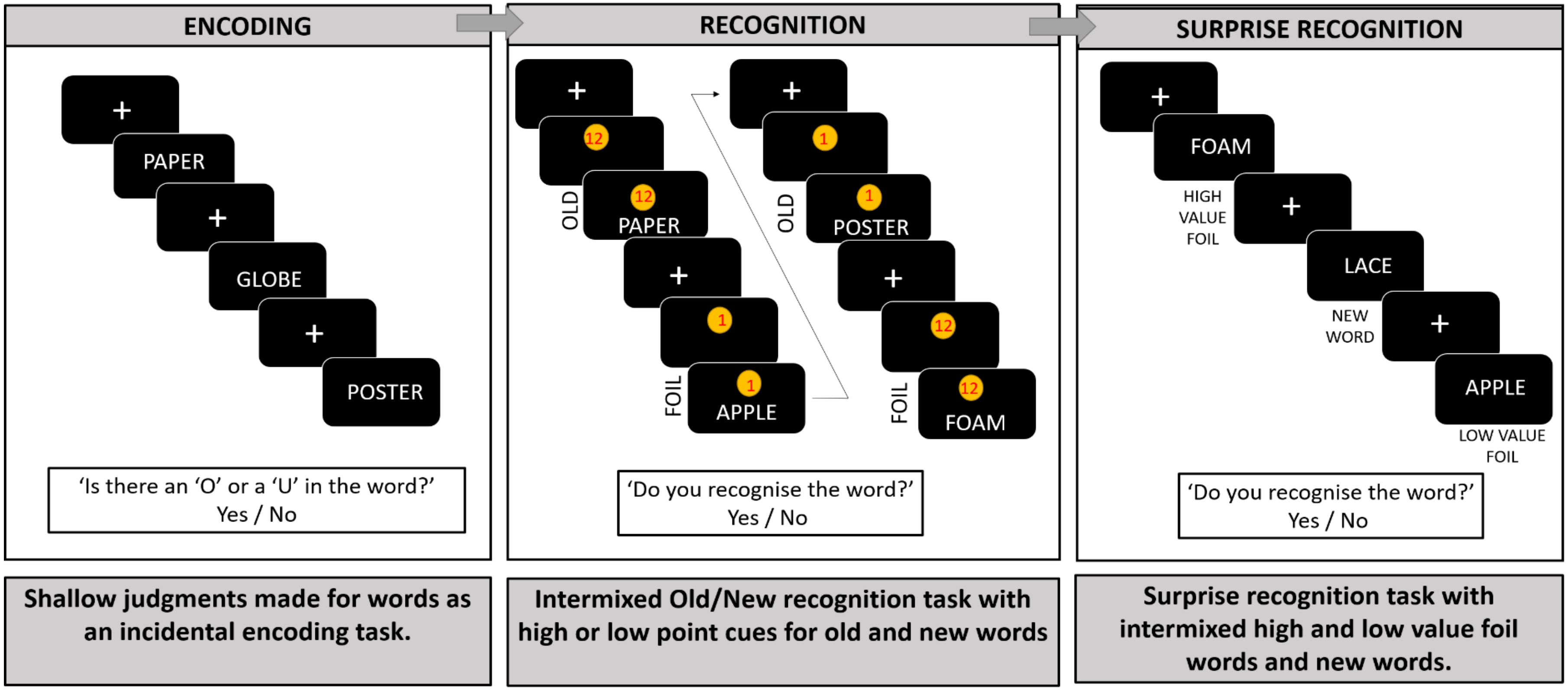


Figure 3.1. Schematic of Experiment 2 procedures. Every participant took part in all stages. During the recognition phase OLD and FOIL words were randomly intermixed and given either a high (12 points) or low (1 point) point value.

3.3.4 Data Analysis

For Recognition Test 1, the proportion of correct responses for each point value (High-12 points, Low-1 point) was calculated separately for Old words (i.e. hit rates) and New words (i.e. correct rejection rates). Mean reaction times (RTs) were also calculated for Old and New words split by test point value. The raw accuracy scores above were used to derive estimates of discrimination ability and response biases during each recognition test, in order to provide more meaningful measures of recognition performance for testing potential differences between the value conditions. Discrimination ability (*Pr)* and Response bis (*Br)* were calculated for Test 1 in an identical way to Chapter 2 (see 2.3.4 for details).

On the Surprise foil test, similarly to Chapter 2, proportion accurate responses were calculated separately for each foil condition (High-12 points, Low-1 point), and for New items (resulting in hit rates for foils since these items are "old" on the final test, and correct Rejection rates for New items). Mean RTs were also calculated for these conditions. As there was only one set of New items and all Foils were intermixed, one measure of Surprise test response bias (*Br*) was calculated for each participant, using a summary measure of *Pr* that was not split according to foil condition (*Br* = false alarm / (1- summary *Pr*). Condition-specific *Pr* scores were then calculated for each foil type by subtracting the common false alarm rate from the separate foil hit rates.

Raw proportion accurate measures (hit and correct rejection rates) were analysed, and complemented by *Pr* and *Br* measures. *Pr* was used as the key measure of recognition discrimination ability for inferential statistical analyses of both Test 1 and Surprise test data, in order to make the results comparable across phases and to compensate for individual differences in response biases. The *Br* measure was also included as a key measure to investigate potential point value differences in response biases. Likewise, differences in RTs across conditions were also analysed. For details of the frequentist inferential statistical tests and the Bayesian statistics tests refer to detailed descriptions in Chapter 2 (see section 2.3.4).

3.4 Results

Raw individual participant level data from all recognition tests in all phases can be found in an anonymised format on the Open Science Framework website along with other supplemental materials such as Bayesian analysis output including robustness checks (https://osf.io/k6gqw/?view_only=458043614c7541fc909b18e3645860f7).

3.4.1 Recognition task 1: Point Manipulation

Recognition performance was calculated separately for High and Low Value point words, but was taken from the same recognition test phase as these items were intermixed. Performance on the first recognition test is shown in Table 3.1. As can be seen here, recognition accuracy and discrimination varied only very slightly between High and Low value items. Response bias and reaction times however appeared to be more affected by the value manipulation.

Table 3.1. *Performance on the first recognition test. The table shows proportion accurate responses, discrimination (Pr), response bias (Br) and mean reaction times, split by recognition point value.*

Point Value	Recognition accuracy		Discrimination (Pr)	Response Bias (Br)	Reaction Times (*ms*)	
	Old Words	New Words			Old Words	New Words
High Value	.57 (.17)	.72 (.19)	.28 (.21)	.39 (.20)	892 (125)	901 (143)
Low Value	.55 (.16)	.74 (.18)	.28 (.19)	.36 (.18)	876 (139)	885 (140)

Note: Values shown are Means (SD). *N*= 86

A repeated measure ANOVA with Point Value (High Value v Low Value) and Word Type (Old v New word) showed no main effect of Point Value on recognition accuracy (*F*(1,85)= 0.03, *p*=.863, n_p^2< .001; *EMM*; High Value=0.64, *SE*=0.01; Low Value= 0.64, *SE*=0.01) but a main effect of Word Type (*F*(1,85)= 33.39, *p*<.001, n_p^2= .282; *EMM*; Old word=0.56, *SE*=0.02; New word= 0.73, *SE*=0.02), with New words being recognised significantly more accurately than Old words. There was also a significant interaction between Point Value and Word Type (*F*(1,85)= 6.67, *p*=.012, n_p^2= .073), driven by opposite effects of the value manipulation on New and Old words. Whereas accuracy for Old Words increased when those words had High rather than Low value (*t*(85)= 1.82, *p*=.072, *d*= 0.15; BF_{10},= 1/1.736), accuracy for New words decreased when those words had High compared to Low value (*t*(85)= 2.22, *p*=.029, *d*= 0.12; BF_{10},= 1.207), however these differences should be tentatively interpreted due to both supported by only anecdotal Bayes Factors.

Because raw measures of proportion accurate responses on old/new recognition tasks can be influenced by both memory and response biases, I also calculated a measure of discrimination (*Pr*) to estimate participant's ability to discriminate between old and new items while adjusting for possible response biases. This showed that even though there was an interaction between Point Value and Word Type for raw accuracy, this was not accompanied by a change in discrimination ability. Overall there was equal discrimination (*Pr*) between the High and the Low Value items (Figure 3.2) with no difference in discrimination between High Value and Low Value words (*t*(85)= 0.17, *p*=.863, *d*= 0.01; BF_{10}= 1/8.279).

I then analysed the *Br* measure of response bias to estimate participants' tendency to guess 'old' when uncertain (Table 3.1). Participants showed a laxer response bias (i.e. an increased tendency to guess 'old') for the High Value compared to the Low Value words (*t*(85)= 2.56, *p*=.012, *d*= 0.16; BF_{10}= 2.57) which was significant but had a small effect size, as seen in Figure 3.2 and was supported by anecdotal Bayesian effects. This change in response bias therefore led to higher accuracy for Old Words but lower accuracy for New words in the High compared to Low value conditions

For Reaction Times, a repeated-measures ANOVA with Point Value (High vs. Low Value) and Word Type (Old vs. New word) indicated a main effect of Point Value on Reaction times (*F*(1,85)= 6.37, *p*=.013, n_p^2=*.07*; *EMM*; High =896, *SE*=14; Low=881 , *SE*=14) with slower responses to High Value words than Low Value words. There was no main effect of Word Type (*F*(1,85)= 1.51, *p*=.222, n_p^2=*.02*; *EMM*; Old word =884, *SE*=14; New word=893 , *SE*=15) and no interaction between the factors (*F*(1,85)<.001, *p*=.996, n_p^2<.001).

Thus, performance on the first test showed that participants were not able to increase their discrimination performance for higher value items, but they did change both their speed of responding and their response bias, suggesting that the value manipulation was successful at changing retrieval processing.

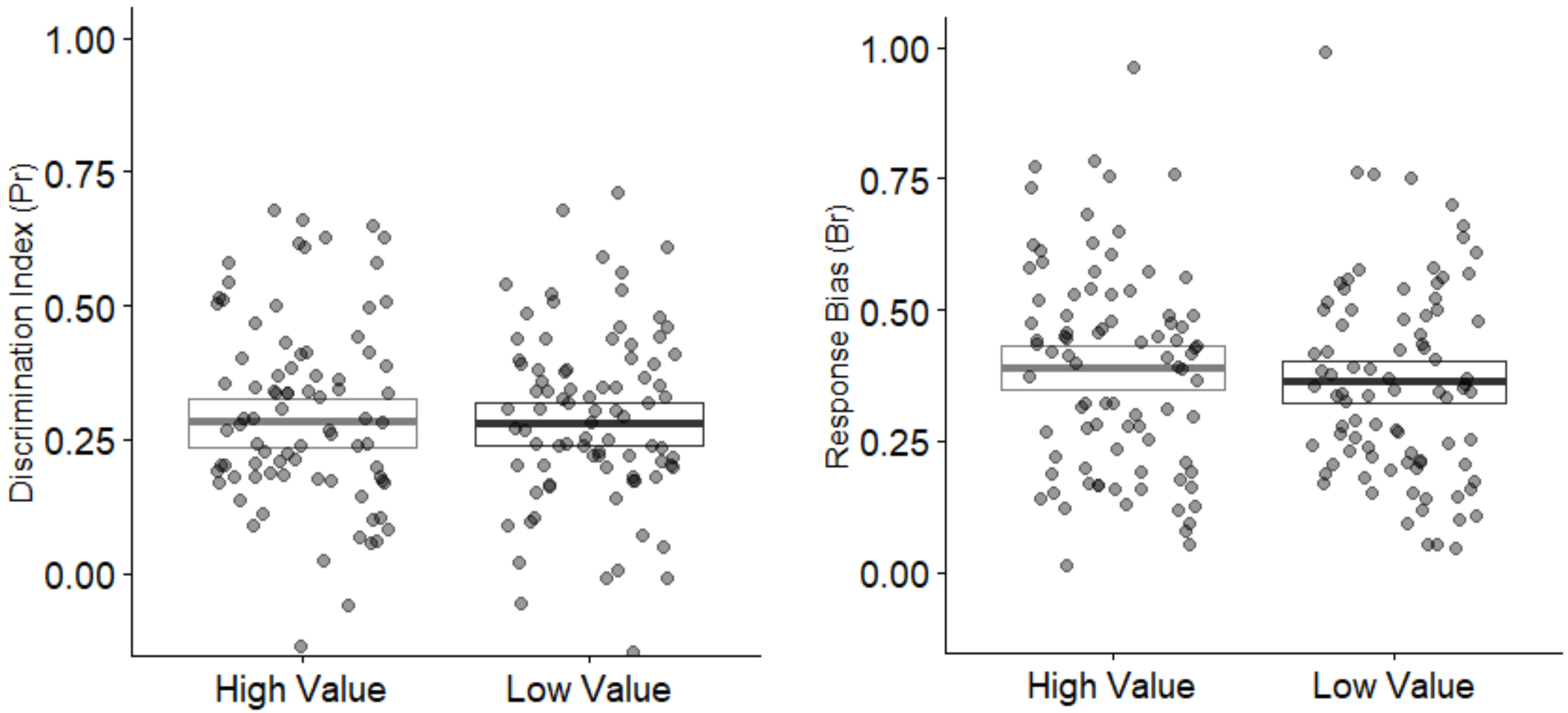


Figure 3.2. Discrimination index (*Pr*) (left) and Response Bias (*Br*) (right) for High and Low Value words during the recognition test 1. The dot plot shows each individual score, with the mean value and confidence interval plotted as a centre line and box respectively. Scores have been randomly scattered across the x-axis for visualisation purposes.

3.4.2 Surprise Test: Subsequent Foil Recognition

On the surprise foil test, New word identification was relatively high (*M*= 0.75, *SD*=0.15) and fast (*M*= 781ms, *SD*=174ms). The overall response bias score was calculated for all foil words and New words (Overall *Br* = false alarm / (1- summary *Pr*. See 3.3.4 for the calculation of an 'overall *Br* score') was close to neutral (*M*= 0.47, *SD*=0.20).

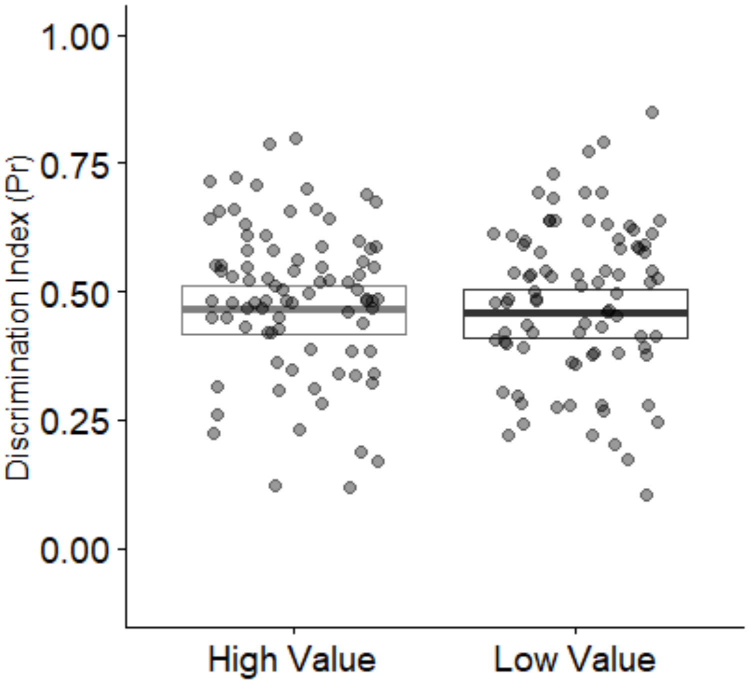


Figure 3.3. Pr discrimination for High and Low Value words during the surprise foil test. The dot plot shows each individual score, with the mean value and confidence interval plotted as a centre line and box respectively. Scores have been randomly scattered across the x-axis for visualisation purposes.

Next, *Pr* was compared across the two foil conditions to investigate whether the points manipulation during the first test had influenced foil encoding[2]. Foil word discrimination (*Pr*) did not differ depending on whether foils previously had a High Value (*M*=0.46, *SD*=0.22)

---

[2] Since all types of foils were intermixed in the final test, any differences between foil conditions in accuracy (hit rates) would be the same in terms of the discrimination measure *Pr* (*Pr* is calculated here by subtracting false alarms to new items from hits to Old items, and with only one category of New items this results in subtracting a constant from all old conditions; Snodgrass & Corwin, 1988).

compared to a Low Value (*M*=0.46, *SD*=0.23; *t*(85)= 0.90, *p*=.373, *d*= 0.03; BF_{10}= 1/5.709). Reaction times on the final test also did not vary between High Value foils (*M*= 733ms, *SD*=153) and Low Value foils (*M*= 723ms, *SD*=145; t(85)= 1.23, *p*=.224, *d*=0.07; BF_{10}= 1/4.084).

3.4.3 Self-reports

The difference between each participant's High Value and Low Value recognition discrimination scores (Pr on Test 1) were calculated and correlated with self-report measures of motivation, effort and difficulty, to investigate if individual differences in these variables were related (see Table 3.2 for abbreviated questions asked).

Table 3.2. *Post-Experimental Questionnaire.*

Question	Post Experiment Questionnaire	Response Options
1	Did you use any techniques or strategies to help you learn the words?	Y/N
2	Did you use any techniques or strategies to help you recognise the words in the first test?	Y/N
3	How motivated were you to win the maximum amount of points during the first test?	scale 1 to 5
4	How difficult did you feel the first test was to recognise the words you had earlier judged?	scale 1 to 5
5	How much effort did you put into recognising the words in the first test and winning the maximum amount of points during the first test?	scale 1 to 5
6	Did you anticipate the final surprise test at the end?	Y/N

Note: Y/N means Yes/No response options, For the scale of 1 to 5: 1 represents low endorsement of the item, with 5 being the highest level of endorsement.

There were no relationships between the difference in Point Value discrimination ability (High *Pr* – Low *Pr*) or Response Bias (High *Br* – Low *Br*) with Motivation, Effort or Rask Difficulty as seen in Table 3.3.

Table 3.3. *Spearman's correlations between the difference in self-report measures of motivation and effort between reward types with measures of improvement on recognition scores during test phase 1 and the surprise test.*

Self-report measures	Test 1 *Pr*			Test 1 *Br*		
	r_s	*p*	BF_{10}	r_s	*p*	BF_{10}
Motivation	-.202	.062	1/1.108	-.103	.344	1/3.939
Effort	-.174	.109	1/1.077	.048	.661	1/7.327
Difficulty	-.127	.243	1/3.477	.026	.810	1/6.876

Note: All correlated variables are calculated based on a difference measure between High Reward and Low reward scores. (i.e. Motivation at study = Motivation during the High Reward study phase minus Motivation during the Low Reward study phase.) Bayesian correlations calculated with a Pearson's r width of 1). All correlations are based on *N*=86.

3.5 Discussion

The current experiment examined the effects of reward, in the form of point values, on retrieval, with the aim to manipulate incidental encoding through prioritised retrieval processing for high compared to low value items. This study extended on previous findings using the value-directed remembering paradigm where high value items are strategically prioritised over low value items (Castel et al., 2002, 2007; Hennessee et al., 2017). Previous research with the value-directed remembering paradigm has focused on the effects of value on intentional encoding as tested with free recall, however this novel study examined the effects of value on retrieval processing during a recognition task. The main aim was to see if value influenced incidental encoding during retrieval attempts, similarly to how value affects incidental encoding of non-target items when taking part in other judgment tasks incidental value (Adcock et al., 2006; Mather & Schoeke, 2011; Wimmer, Braun, Daw, & Shohamy, 2014). Value was seen previously to increase incidental encoding of items through automatically increasing motivational and anticipatory reward-seeking behaviour through influencing attention

(Anderson, 2016; Miendlarzewska et al., 2016) and monitoring processes (Chiew & Braver, 2016). Importantly, the effect of value on incidental encoding during recognition itself has not been explored before, and this study bridges a gap between the well documented intentional, strategic effects of value on intentional encoding (Castel et al., 2007, 2011, 2013; Cohen et al., 2014; Hennessee et al., 2017, 2019) with the wider literature on factors that influence incidental encoding of foils during recognition (Bergström et al., 2015; Gray & Gallo, 2015; Jacoby, Shimizu, Daniels, et al., 2005; Vogelsang et al., 2016, 2018). Implementing the manipulation of reward during retrieval enabled both response bias and discrimination to be examined, something that cannot be done if rewards are manipulated at encoding only.

In this experiment however, I found that varying levels of value during a recognition task did not have an effect on retrieval success (recognition test 1). Similarly, there was no effect of value on incidental encoding of foils as indicated by subsequent recognition rates. It may be the case that rewards are more likely to alter encoding strategies more directly, rather than have a marked effect on strategic retrieval, to the extent that there is a shift in incidental encoding rates. Interestingly however, value did have an effect on response bias during the first test, with a shift in response bias being more conservative on low value trials to less conservative on high value trials. In other words, for high value words, participants had a more neutral response bias when uncertain, compared to for low value items, where they were more likely to guess ‘new’ when they were uncertain or guessing. This finding suggests that there was, to some level, a change in retrieval processing, with participants being less likely to dismiss high value items when guessing as being ‘new’. This finding is further supported by response time differences between high and low value items. Participants had longer response times for high value over low value words, irrespective of word type (old or new). This suggests that there could have been increased monitoring or processing of the high-value words in line with the suggestion that reward cuing close to the time of stimuli presentation can lead to reactive control processes (Chiew & Braver, 2016).

A small set of studies have similarly explored the effects of reward on retrieval, with differing study aims not centred around incidental encoding during recognition itself (Bowen, Marchesi, & Kensinger, 2020; Han et al., 2010; Marini, Marzi, & Viggiano, 2011; Yan et al., 2018). But informatively, on the whole, these studies similarly found altered response biases, confidence ratings or response times. Generally, they found a less conservative bias when responding to high value items and shorted response times. My findings converse with these prior studies, similarly finding a less conservative bias for high value item. However, response times seem to differ, with here higher value items taken longer to respond too. This could be due to differing task demands having an effect on response time (Chiew & Braver, 2016). If retrieval processing was altered due to reward value, for example through automatic attention prioritisation, and more internally driven effort and increased elaborate deliberation and item processing, then this would be more cognitively taxing than processing low value items as these were less important, holding only a twelfth of the high reward value. Likewise, to this chapter's findings, these studies found either no effect of reward on retrieval accuracy (no improvement to hit rates), or an improvement in accuracy that was explained fully by changes in response bias criterion. This is interesting and poses questions about the effectiveness for rewards to actually alter recognition performance and ability.

My results are also consistent with more sensitive measures such as fMRI (Han et al., 2010) and ERPs (Halsband et al., 2012), which have shown neural differences in recognition processing due to reward value with small or limited behavioural difference. These studies suggested that the effects of rewards on processing are too small to detect behaviourally, but showed neural differences where rewards altered brain processes through mechanisms such as motivation related to goal-attainment (Han et al., 2010), and the reinstatement of non-strategic reward-related memory processes (Halsband et al., 2012). Thus, even though there is limited evidence of the effects of reward on recognition processing, my results seem to converge with the existing literature that people alter their retrieval processing in response to rewards, but such changes may be rather subtle and difficult to detect behaviourally.

The limited research indicated that the effect of external rewards on recognition processes are important to assess. Part of the issue with examining changes in recognition processes in prior literature has been rooted in the inability to calculate discrimination indices and response biases. This is related to paradigm constraints when manipulating reward at encoding. Primarily, these paradigms have been focussed on the effect of reward on encoding showing an increase in encoding relating to the anticipation or association of rewards or higher values (Adcock et al., 2006; Cohen et al., 2017, 2014; Gruber et al., 2013; Hennessee et al., 2019; Mather & Schoeke, 2011). The overarching issue is, however, if an 'old/new' recognition task is used, evaluating only 'hit' rates can be misleading, with the risk of inflated the effect of reward on recognition (Donaldson, 1996). Importantly when looking at recognition scores, it is standard practise to look at both correct 'old' responses and correct 'new' responses (Snodgrass & Corwin, 1988). Paradigms like the one used in this experiment enables the examination of discrimination and response bias through equally weighting high and low rewards to old and new items. This then starts participants off with a bias that should in principle be neutral. There is no incentive for participants to un-proportionally guess 'old' or 'new' for either the high or low value items, whereas they are rewarded for overall accuracy in recognising old items and identifying new items. This is something that was stressed to the participants in the instructions.

A practical change to increase the potency of the reward value in light of subtle reward-based changes, could be that the points directly correspond to monetary winnings. In this experiment it was not possible to assess individual motivation scores separately for high and low value items as the test was intermixed, so would have been unrepresentative of their actual motivation levels. However self-reports of participants motivation and effort to win the most points were collected. Here no relationship was seen between motivation or effort scores with the difference in discrimination ability. Though these findings are limited, it does indicate that motivation during the task could be boosted further. In this current experiment, mainly due to financial limitations, the points were entered into a league table where the top three scorers won. This may have been highly motivating for some participants, but others might have felt that being among the top three is unlikely, and this could have reduced the motivational effect intended by the points. Having a direct link between each participant's accuracy to a tangible monetary value could have ensured that all participants were highly motivated to take part.

3.6 Conclusions

Overall, the findings from this experiment add to our understanding about the ability for value to alter recognition performance and incidental encoding during recognition attempts, which has not been previously examined. This novel experiment, which combined the memory-for-foils paradigm with the value-directed remembering paradigm, has allowed me to examine how value affects strategic retrieval processing, and if such altered processing affects subsequent recognition of foils. Overall, recognition decision speed and response biases changed due to value, but with discrimination performance unaltered and no change in incidental foil encoding. These findings provide insight into the limits of value prioritisation on memory performance. It supports previous conclusions that value-directed remembering manipulations involve motivational shifts and works through altering prioritisation of strategy use and attentional resource allocation (Castel, 2007; Castel et al., 2007; Cohen et al., 2014; Hennessee et al., 2019; Middlebrooks & Castel, 2018; Middlebrooks, Murayama, & Castel, 2016), yet adds to these conclusions by showing that value prioritisation has more limited effects on recognition attempts than on initial encoding. Importantly, there is an interesting contribution of the effects of value on response bias, something currently that has limited evidence for (Bowen et al., 2020; Han et al., 2010; Marini et al., 2011; Yan et al., 2018). This provides some insight into the ability of value to alter recognition processing, something that is difficult to examine in the wider literature that looks at how value alters encoding rates (*cf.* Miendlarzewska, Bavelier, & Schwartz, 2016).

3.7 Summary

This chapter explored the effect of value on recognition during retrieval attempts. These findings suggest that value can alter recognition processing in terms of altering one's bias, yet value did not alter recognition performance or incidental encoding rates. This experiment provides valuable contributions to the limited evidence on how value alters recognition processing, adding to the wider literature on value's effect on encoding and memory more broadly (Arias-Carrion et al., 2010; Hidi, 2016; Miendlarzewska et al., 2016). In the next chapter, I present further experiments that extended on the findings in both Chapter 2 which found effects of value during an intentional encoding task on subsequent discrimination ability, and the findings in this chapter which found that value altered strategic retrieval processing, but not discrimination ability. In the next experiment, I manipulated value during an intentional encoding task and modified the subsequent recognition task to explore how value interacts with subjective experiences of recognition processes during retrieval attempts, and how these factors interact to affect incidental encoding of foils.

Chapter 4: Effects of reward and subjective experience of recognition on incidental encoding of novel items during retrieval attempts.

4.1 Overview

In this chapter I extended on the previous chapters to provide further insight into the effects of reward on memory encoding and retrieval. In the previous two experiments, rewards affected motivational states and information processing but were seen to only alter intentional encoding in Experiment 1, through strategic prioritisation of highly valuable information in young adults. Rewards also affected retrieval speed and response biases in Experiment 2, which may reflect modulations of recognition processes related to memory judgments and monitoring. This raises questions about how value interacts with recognition judgements and subjective experiences of recognition such as memory confidence and updating of memory after corrective feedback which the next chapter assesses.

4.2 Experiment 3

The anticipation of highly valuable or rewarding information leads to increases in encoding and subsequent retrieval of that information (Adcock et al., 2006; Spaniol et al., 2014). As discussed in previous chapters, explanations of reward enhancement on memory vary, with research suggesting immediate memory benefits are due to increased motivation and prioritised attentional or strategic resources to rewarding information (Castel et al., 2007; Cohen et al., 2014; Hennessee et al., 2017; Middlebrooks, 2018), compared to automatic enhancements which may rely on consolidation (Miendlarzewska et al., 2016; Murayama & Kitagami, 2014). Previously in Experiment 1, I observed a reward-related modulation of intentional encoding for young adults as they reduced the amount of non-rewarded information they learned. Again, in Experiment 2, there was a shift in response bias between high and low point value words when value was assigned during a recognition test. These findings, taken together with the existing literature suggests that there is an effect of reward on intentional memory processes (Castel, Benjamin, Craik, & Watkins, 2002; Castel et al., 2007, 2013; Cohen et al., 2014; Hennessee et al., 2017; Middlebrooks & Castel, 2018). Information that holds higher value or importance is monitored more efficiently (McDonough, Bui, Friedman, & Castel, 2015) and is reported to be more vivid (Gruber & Otten, 2010).

Recognition processes and in particular subjective experiences of recognition and memory monitoring play a major role during encoding (Finley & Benjamin, 2012) and retrieval (Finley et al., 2010). These processes are important because they enable successful use of memory to guide decision making in a variety of everyday situations, including educational settings. Subjective experiences of recognition can provide insights about recognition success, commonly high confidence is associated with increased likelihood that the judgement will be accurate (Brainerd, Reyna, Wright, & Mojardin, 2003; Koriat & Goldsmith, 1996). However, higher confidence does not necessarily indicate increased memory accuracy (Ingram et al., 2012; Wixted & Mickes, 2010). Furthermore, confidence judgments can influence memory modification and updating through processes like post-retrieval monitoring (Finn, 2017) and feedback (Metcalfe & Finn, 2011). It is therefore important to understand how subjective experiences of recognition interacts with external factors that alter motivation, such as reward.

As discussed previously, the salience of information plays a role in encoding success. Highly salient information attracts attention and is encoded into memory (Anderson, 2016). Reward anticipation or experience is one factor that influences salience, however other sources of salience also influence memory, such as novelty (Krebs, Schott, Schütze, & Düzel, 2009; Wittmann, Bunzeck, Dolan, & Düzel, 2007), or unusual or surprising information (Fazio & Marsh, 2009). Surprising feedback is highly salient and leads to increased automatic memory updating (Butterfield & Metcalfe, 2006; Fazio & Marsh, 2009). The 'hypercorrection effect' demonstrates this, with high confidence memory errors being corrected more than low confidence errors, when participants are given feedback that shows the correct answer (Butler, Fazio, & Marsh, 2011; Butterfield & Metcalfe, 2006; Metcalfe & Finn, 2011; Mullet & Marsh, 2016). The hypercorrection effect has mainly been studied within the context of updating of semantic memory; however, the effect of feedback more generally has a similar effect on episodic memory encoding (Mather & Schoeke, 2011).

To my knowledge, previous literature has not addressed how confidence and feedback during retrieval attempts affects incidental encoding of novel items in an episodic recognition task, and whether this interacts with other salient information such as memory value. Therefore, the aims of Experiment 3 were to examine if incidental encoding during recognition attempts was influenced by subjective experiences of recognition, such as those involved in accuracy/confidence monitoring and feedback-related error monitoring and if these interacted with rewards.

As in the previous two experiments (Experiment 1 and 2), the memory-for-foils paradigm (Jacoby, Shimizu, Daniels, et al., 2005) was modified to enable the examination of both intentional and incidental encoding. As in Experiment 1, a reward manipulation was conducted across two blocks of intentional encoding and initial recognition tests, and these initial encoding-retrieval blocks were followed by a surprise recognition test for foils from the two prior tests. During the initial tests, participants also reported their confidence in each recognition decision, and were given feedback as to whether their decision was correct or not. This enabled me to investigate whether confidence and accuracy during an initial recognition judgement for new foil items (i.e. items that should be given a “correct rejection” response) on a test predicted whether those items would later be recognised in a subsequent surprise foil test. Furthermore, the feedback was added to elicit salient, corrective information based on the reasoning that feedback may enhance encoding of foils when a participant’s memory expectations were contradicted (i.e. eliciting prediction errors) and to investigate if rewards enhanced the effects on encoding of such memory expectancy violations.

I also modified the subsequent foil recognition test, to investigate whether subsequent recognition of foils was associated with different recollection or familiarity experiences depending on prior subjective recognition experience and reward processing. To address the latter issue, I used the Remember-Know judgment scale (Donaldson, 1996; Gardiner & Richardson-Klavehn, 2000) to provide qualitative information about differences between subsequent memory types (recollection and familiarity), which are not distinguishable with an 'Old/New' recognition judgment alone (Migo, Mayes, & Montaldi, 2012).

This new design therefore enabled me to examine potential interactions between rewards, recognition judgement accuracy and confidence, in terms of how these factors affect incidental encoding during retrieval attempts. I predicted that high value rewards would enhance intentional encoding (Adcock et al., 2006; and as found in the young group in Experiment 1) and potentially to also enhance incidental encoding of foils (Mather & Schoeke, 2011). Specifically, I expected that incidental encoding of high value foils, that were encountered during a recognition test, would result in enhanced subsequent recollection on the surprise test compared to low value foils (Wittmann et al. 2005; although see Experiments 1 and 2 for a failure to find reward effects on foil encoding). I also expected that high value foils on the surprise test would have a different memory experience, potentially leading to more 'Recollect' responses than 'Familiar' responses compared to low value foils. I anticipated that confidence and accuracy during initial recognition judgements would have an effect on incidental encoding of novel foils. I further expected that novel foils that were recognised incorrectly as previously seen (i.e. given a "false alarm" response) would be more likely to be later accurately recognised in the subsequent test due to enhanced encoding as a result of surprising error feedback (Metcalfe, 2017; Mullet & Marsh, 2016; Pashler, Cepeda, Wixted, &

Rohrer, 2005). Such enhanced foil encoding in response to error feedback might also be particularly pronounced in the high reward compared to low reward block, if these factors interact to influence incidental encoding during retrieval attempts.

4.3 Methods

4.3.1 Participants

Results are based on 48 native English-speaking student participants (12 Male; *M* age=19.45 years, *SD*=2.16, range=18-31). This sample size (same as originally planned in Experiment 1) enabled me to achieve >0.9 power to detect an effect size of Cohen's *d*=0.7 at two-tailed α=.05 (and to fully rotate and cross all counterbalancing factors, task orders, stimuli assignment to conditions, etc.). The participants were recruited from the University of Kent Psychology participation scheme where students received course credits for their time. Participants had normal or corrected to normal vison (colour-blindness was not tested and only self-reported), and reported not having dyslexia or diabetes, not suffering from disorders like depression and anxiety, and not taking any psychoactive medication. Participants could have been left or right handed. Lastly, participants were only included if they liked chocolate (reported on average liking chocolate 7.8 out of 10, with 10 being that they loved chocolate), and lactose-free chocolate was substituted for milk chocolate if the participant was lactose intolerant. Importantly, participants had not taken part in any other experiments within this thesis.

All participants provided full written consent and the procedures were approved by the University of Kent School of Psychology Ethics Board. Importantly, participants were not aware of the chance to win chocolate before they signed up for the study. Participants were tested either by *LS* or by a master's student who was trained by LS. All testing was conducted in a face-to-face manner.

4.3.2 Materials

The experiment was programmed and run in PsychoPy (Peirce, 2007). A total of 384 English words were used as stimuli (3 - 11 characters, Kucera-Francis frequency range 0-591), of which 75% had been used in Experiment 1, taking an additional 96 words from the words from a larger set used by Vogelsang et al. (2016, 2018), which were originally sampled from the MRC psycholinguistic database (Wilson, 1988). Stimuli numbers were increased from Experiment 1 to provide more power for within-participant analysis. The words were randomly sorted into six lists with 64 words in each, with assignment of lists to conditions counterbalanced across participants, as was reward block order. This yielded a full counterbalance rotation of 12 participants. No counterbalancing of handedness was included as behavioural results of accuracy were the primary measure, differences in RT between 'Old' and 'New' responses were not of interest here, and the interaction with reward could still be examined.

Food-based external rewards were used in this experiment. Milk chocolate buttons were selected as the food-based reward due to the suitable portion size which was easy to conceptualize in terms of trial-by-trial reward winnings. The reward used varied from the earlier chapters (Experiments 1 and 2) where monetary based rewards were used. The decision to move from monetary rewards to food-based rewards were for two reasons. Firstly, food based 'primary' rewards have been seen to have more of an automatic effect on cognition (Beck, Locke, Savine, Jimura, & Braver, 2010) compared to secondary rewards such as money or points. Food based rewards were therefore used as I aimed within this experiment to enhance the incidental, automatic effects of reward on encoding. Secondly, feasibility is a factor, and food-based rewards are easier to obtain for larger samples as they are not as costly as direct monetary rewards. Participants were told that the maximum amount of chocolate winnings they could receive was three 'treat size' bags, equivalent to 45g of chocolate. Participants could take the chocolate winnings away and did not have to consume any chocolate during the study session.

As in the previous two chapters, a set of 160 emotionally neutral images from the IAPS database (Lang et al., 2005) were used in a visual search filler task. This was a 50% subset of the IAPS images used in Experiment 1 (Chapter 2). Apart from this, the filler task was the same as described in Experiment 1 except there were only two repetitions of the task presented, once after the encoding and then again after the retrieval task.

4.3.3 Design and Procedure

The procedure was modified from a memory-for-foils paradigm (Jacoby, Shimizu, Daniels, et al., 2005; Vogelsang et al., 2016) and expanded on the Reward Motivation manipulation design used within Experiment 1. As with the last two experiments, all participants first completed encoding tasks, where in this experiment reward value was manipulated across blocks. Then participants completed separate recognition tasks for each reward value and finally a surprise foil recognition task for high and low value foils as seen in Figure 4.1. The order of reward value presentation was counterbalanced across participants to account for order effects. Again, like in previous chapters, in between each experimental stage a visual search filler task was used.

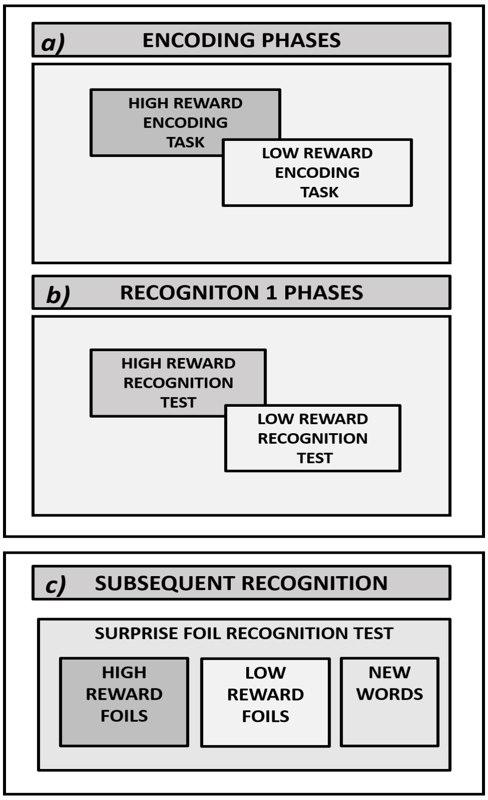


Figure 4.1. Schematic of Experiment 3 phases. Every participant took part in all stages; A) both the High and Low Reward encoding phases, followed by b) both High and Low Reward recognition phases and finally, c) the Surprise Test. Reward blocks were counterbalanced for order. On the Surprise test all foils were intermixed equally with new words.

4.3.3.1 Main recognition task

First, the participants took part in two intentional encoding tasks, one with 'High Value' where they could win more chocolate, and one with 'Low Value' where they could win less chocolate. The order of these tasks was counterbalanced across participants. In each encoding task, participants were presented with a list of 64 words and told to try to learn as many words they could to maximise their chocolate winnings on the recognition tasks. Each word was shown for 2200ms and preceded by a fixation cross for 700ms, the words within each list were presented in a random order. Before each study phase, a reward cue was shown that indicated the amount of chocolate, they could win for learning words; either three chocolate buttons (High Value) or a quarter of a chocolate button (Low Value) per word. Participants did not have to give responses during the intentional encoding phase, but were simply told to try their best to memorise the words for a later test (without being given any specific learning instructions).

Participants then took part two separate recognition tests where each test contained 64 old words randomly intermixed with 64 new foil words. The recognition tests were conducted separately for the High Value and Low Value items, and were presented in the order of encoding (so if the High Value encoding task was conducted before Low Value encoding, then the High Value test was conducted before the Low Value test, and vice versa). Participants were explicitly told which test phases corresponded to High or Low Value conditions and informed that old items would only be drawn from the corresponding High or Low Value encoding task. This information was given verbally and with a visual cue and written text at the start of each test phase, to ensure participants were fully aware of which test corresponded to the reward value (either 3 buttons or ¼ of a button per correct response). Participants were instructed that they would be winning the corresponding chocolate value for all correct responses, both for recognising studied words and for correctly identifying new words, and that the higher their accuracy during the recognition test the more they could win.

Each test trial began with a 700ms fixation cross, after which participants were shown the word for 2200ms before being able to respond. Next, the response options ('old' and 'new') appeared on the screen and participants then had 2700ms within which time to respond. As with the previous experiments, participants used the 'Z' key to respond 'Old' with their left-hand index finger or the 'M' key to respond 'New' with their right-hand index finger. Participants were explicitly instructed not to respond until the response options appeared on the screen.

Along with providing 'Old/New' recognition judgments about the words like in previous experiments, participants additionally provided confidence judgments about their memory decision. Confidence was measured by using a continuous scale involving holding down the keys for different durations to indicate confidence (described below, based on Bergstrom et al., 2015). When participants gave their ‘Old’ or ‘New’, responses (‘Z’ and ‘M’ keys – first key response time), they were asked to simultaneously indicate their confidence in their decision by holding down the keys either longer or shorter (duration of key press - key release time), and they were given visual feedback on the screen by colour changes to the chosen response option to help them calibrate their confidence judgement. Specifically, participants were told that short button press durations would indicate an unconfident memory decision, matched by a small colour change to their chosen ‘Old/New’ judgment, whereas a longer button press indicated higher confidence in their decision and larger colour changes. The colour change was controlled by button press duration which changed the colour in small intervals on the RGB (Red-Green-Blue) colour scale; from bright green indicating not confident at all to bright red indicating highly confident (and colours in between indicating different degrees of confidence). This setting resulted in a gradual colour change to the ‘Old/New’ test options on the screen, with the time to make a full colour change 1659ms, meaning that participants needed to respond within 1000ms of the response options appearing on the screen in order to have time to also use the full confidence scale (based on piloting, it was established that this was possible on the vast majority of trials since participants had already had 2200ms to view the word before the options appeared – see above).

After each recognition trial, participant’s received visual feedback indicating if their response was correct or incorrect. The feedback was shown by a white circle appearing around the correct response, indicating if the participant was correct (congruent response-feedback) or incorrect (incongruent response-feedback). Feedback was shown for 1300ms. Thus, each trial lasted 6900ms in total (700ms fixation, 2200ms pre-response word presentation, 2700ms response window, finishing with 1300ms post response feedback; see figure 4.2 for schematic of trial timings). Any trials that were not responded to were excluded from the analysis. Before completing the tests, participants received comprehensive verbal and written instructions and completed a practise session into how to respond and use the continuous measure of confidence and what the feedback circle indicated.

After completing both initial recognition tests phases (High Value and Low Value separately tested) participants were presented with a final surprise recognition task. During the surprise test, participants were informed that any words that they have previously seen during the whole study could appear and these would be presented randomly intermixed with completely new words. In total 256 words were presented in a random order, of which 64 foil words were from the High Value and 64 foil words from the Low Value reward recognition phases, with the remaining 128 words being completely new.

Instead of making an ‘Old/New’ recognition judgment in this final test, participants were asked to make a ‘remember/ know/new’ judgment on a six-point scale; Recollect (1), Strong feeling of Familiarity (2), Weak feeling of Familiarity (3), Would have to Guess (4), Unsure New (5), Very sure New (6). Instructions for the ‘remember/know’ responses were based on instructions from Gardiner and Richardson-Klavehn (2000). Comprehensive verbal and written instructions about how to respond in the surprise recognition test using the Remember-Know procedure (see Appendix 3.1.2 for task instructions). Participants were given examples of recollected and familiar memories, and asked to produce examples of these to tell the experimenter. The use of the six-point scale was also practised and keyboard sticker references were provided. Participants were instructed not to focus on the speed of their judgment but to focus on the type of memory they experienced and to be as accurate in this decision as possible. Each trial was self-paced, with words presented after a fixation cross for 700ms.

4.3.3.2 Visual Search filler task

In between all experimental phases a visual search task was used as a filler to prevent rehearsal between each experimental task (two times in total). This task was the same as described in Experiment 1[3] (See section 2.3.3.2). Briefly, participants were presented with images and asked to search for a white triangle that had been superimposed onto a subset of images and asked to respond to the question; ‘Is there a white triangle on the image?’ This task lasted for 1 minute in total.

4.3.3.3 Post-Experiment tasks and debrief

After all computerised tasks, participants were asked to complete a questionnaire about their motivation during the study and test phases, the effort they put into learning or recognising words, and if they expected the surprise test, as well as questions about any strategies or memory techniques used. Participants were also asked about how much they liked and ate chocolate. The first 10 participants had a reduced questionnaire, focussing primarily on strategy use. All participants were given verbal and written debrief and were given 3 bags of milk chocolate buttons or lactose-free equivalent (45g). All participants were given the same amount of chocolate winnings.

---

[3] For the first ten participants a letter search task was used, similar to the task described above, but participants responded ‘yes’ if they saw a blue’ X’ in an array of letters (Treisman Letter Search Task). This task was changed to an image based visual search task to prevent the potential of the distractor task interfering with the word recognition consistent with all other experiments. There was no significant effect of distractor task version on word recognition ($F(1,42)=1.18$, $p=.284$, $n^2=.027$) therefore these versions were collapsed.

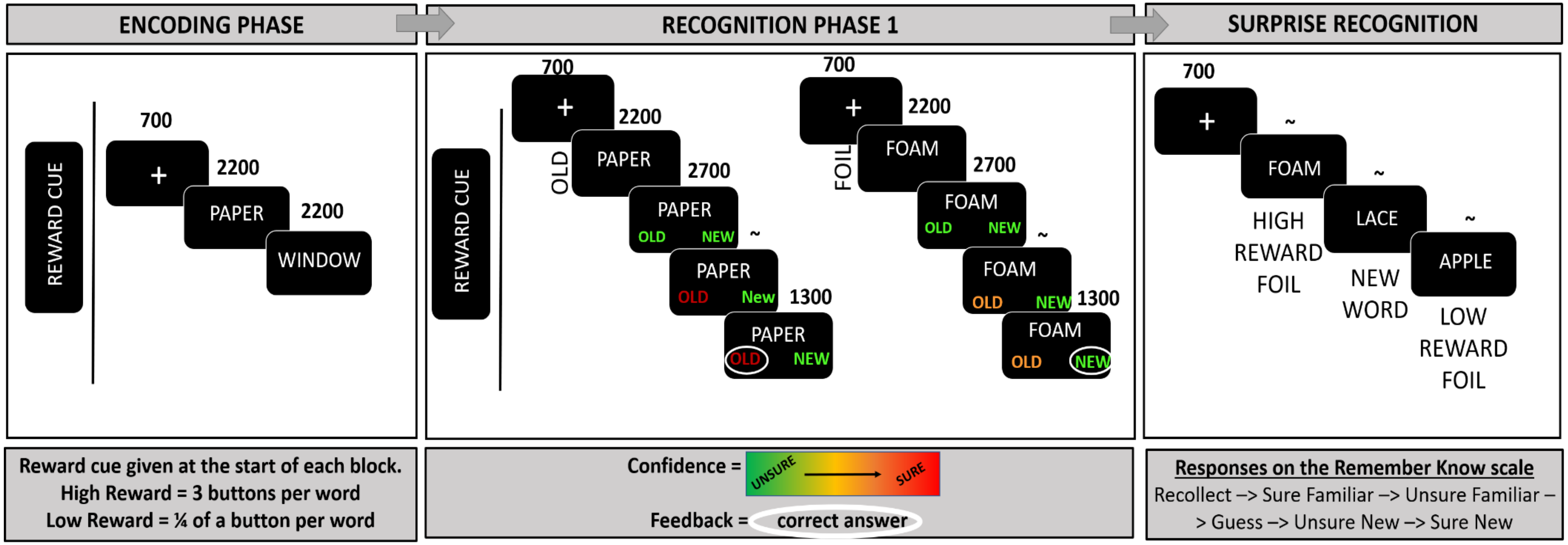


Figure 4.2. Schematic of Experiment 3 trial procedures and timings. Every participant took part in all stages. Reward blocks were counterbalanced for order (High Study-> Low Study->High Test-> Low Test or visa versa with Low reward being presented first). Continuous confidence scores were given during the initial recognition tests and correct answer feedback was provided. On the Surprise test <u>all foils</u> were intermixed equally with new words and Remember- Know judgments were made. All trials had a fixation cross proceeding the next trial. All trial timings are in milliseconds (*ms*) and the tilde symbol (~) indicates self-paced.

4.3.4 Data Analysis

For the Test 1 stage, the proportion of correct responses for each Reward Level (High: 3 buttons, Low: ¼ of a button) was calculated separately for Old words (i.e. hit rates) and New words (i.e. correct rejection rates). Mean reaction times (RTs) and continuous Confidence judgments were also calculated for Old and New words split by reward level. The raw accuracy scores above were used to derive estimates of discrimination ability and response biases during each recognition test, in order to provide more meaningful measures of recognition performance for testing potential differences between the value conditions. Discrimination ability (*Pr)* and Response bias (*Br)* were calculated for Test 1 in an identical way to Chapter 2 (section 2.3.4 for details).

On the Surprise foil test, similarly to in Chapter 2 and 3, the proportion of accurate responses were calculated separately for each foil condition (High: 3 buttons, Low: ¼ of a button), and for New items (resulting in hit rates for foils since these items are "old" on the final test, and Correct Rejection rates for New items). Mean RTs and Confidence was also calculated for these conditions. As there was only one set of New items and all Foils were intermixed, one measure of Surprise test response bias (*Br*) was calculated for each participant, using a summary measure of *Pr* that was not split according to foil condition (*Br* = false alarm / (1- summary *Pr*). Condition-specific *Pr* scores were then calculated for each foil type by subtracting the common false alarm rate from the separate foil hit rates.

Raw proportion accurate measures (hit and correct rejection rates) were analysed, and complemented by *Pr* and *Br* measures. *Pr* was used as the key measure of recognition discrimination ability for inferential statistical analyses of both Test 1 and Surprise test data, in order to make the results comparable across phases and to compensate for individual differences in response biases. The *Br* measure was also included as a key measure to investigate potential Reward Level differences in response biases. Likewise, differences in RTs across conditions were also analysed. For details of the frequentist inferential statistical tests and the Bayesian statistics tests refer to detailed descriptions in Chapter 2 (see 2.3.4).

4.4 Results

Raw individual participant level data from all recognition tests in all phases can be found in an anonymised format on the Open Science Framework website along with other supplemental materials and additional analyses, such as Bayesian analysis output including robustness checks (https://osf.io/va637/?view_only=4f955720696a4dcbaf735c68109e194b).

The effect of reward value on memory will be examined after both intentional encoding (Recognition time 1) and incidental encoding (Subsequent Recognition during the Surprise Recognition Test). I will firstly examine how reward affects not only accuracy after intentional encoding, but also confidence judgments and response times. In terms of subsequent recognition, I will examine how reward affects incidental encoding and any related differences in subsequent memory experience. I will then move on to examine the item-specific, metacognitive-like related factors of confidence, response accuracy and feedback congruency in relation to the effect on incidental encoding (Subsequent Recognition phase). I will examine how accuracy interacts with confidence and reaction times to enlighten me about how these processes interact during retrieval attempts which could lead to differences in subsequent recognition experience and accuracy.

4.4.1 Recognition task 1

Performance on the first recognition tests is shown in Table 4.1. As can be seen here, all measures were rather similar between High and Low Reward Value items and regardless of Word Type. A within-subjects 2x2 (Word Type: Old v New; Reward: High v Low) ANOVA on accuracy (presented in Table 4.2) showed that there was no difference in recognition accuracy due to Reward or Word Type, contrary to predictions. This was somewhat surprising since this paradigm was overall similar to the manipulation of monetary rewards in Experiment 1 (in Chapter 2), which resulted in a statistically significant enhanced accuracy in a Reward recognition test compared to a No-Reward recognition tests after intentional encoding in young adults. The lack of effect of Reward on raw accuracy was echoed in no difference in Old/New Discrimination Ability (*Pr*) between the different levels of Reward Value ($t(47)= 0.19, p=.85, d=0.02$; $\text{BF}_{10}= 1/5.290$) with the mean scores seen in Table 4.1 and individual responses seen in Figure 4.3. In contrast to Experiment 2 (in Chapter 3), where I saw a differences in Response Bias (*Br*) between High and Low Reward Value items, here there was also no difference in Response Bias due to Reward Value ($t(47)= 1.19, p=.239, d=0.21$; $\text{BF}_{10}= 1/4.762$).

Table 4.1. *Performance on the first recognition test for Old and New items. The table shows proportion accurate responses, discrimination (Pr), response bias (Br), mean confidence and mean reaction times, split by High and Low Reward value.*

Reward level	Recognition accuracy		Discrimination (*Pr*)	Response Bias (*Br*)	Reaction Times (*ms*)		Confidence (*ms* duration)	
	Old Words	New Words			Old Words	New Words	Old Words	New Words
High Reward	.74 (.13)	.72 (.17)	.46 (.26)	.50 (.14)	489 (115)	560 (149)	1428 (354)	1121 (338)
Low Reward	.75 (.15)	.72 (.16)	.46 (.23)	.53 (.17)	486 (106)	545 (117)	1423 (351)	1127 (341)

Note: Values shown are Means (SD). Values are calculated based on *N*=48.

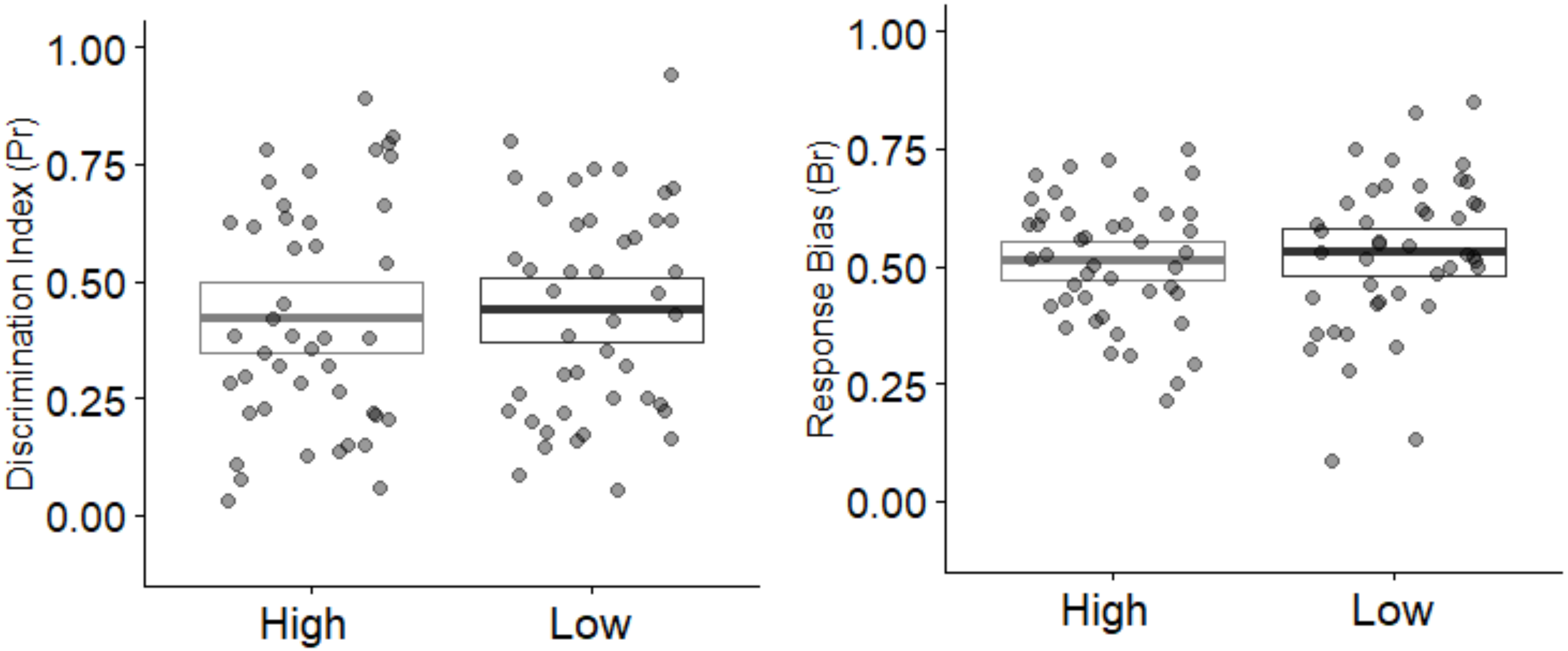


Figure 4.3. Pr discrimination (left) and Br response bias (right) for High and Low Reward words during the recognition test 1. The dot plot shows each individual score, with the mean value and confidence interval plotted as a centre line and box respectively. Scores have been randomly scattered across the x-axis for visualisation purposes. *N*=48.

Table 4.2. *Three separate 2 X 2 ANOVAs testing the effect of Reward (High X Low) and Word Type (Old word X New word) on Recognition Accuracy, Reaction Times and Confidence scores within Recognition Test Phase 1.*

Performance Measure		*F*	*p*	n_p^2
Accuracy	Reward Value	0.035	.852	.001
	Word Type	1.140	.291	.024
	Reward Value * Word Type	0.016	.899	<.001
Reaction Times	Reward Value	0.422	.519	.009
	Word Type	**37.42**	**<.001**	**.443**
	Reward Value * Word Type	0.705	.405	.015
Confidence	Reward Value	<.001	.987	<.001
	Word Type	**91.53**	**<.001**	**.661**
	Reward Value * Word Type	0.134	.716	.003

Note. Significant values below .05 are shown in bold. *N*= 48

On average participants were very fast at responding to 'Old/New' judgments as seen in Table 4.1[4]. After conducting a 2 X 2 ANOVA examining response times as a function of Reward Value (High v Low) and Word Type (Old v New) there was no difference in Response Times due Reward Value as seen in Table 4.2. There were however, differences due to Word Type, with responses to Old Words (*M*= 487, *SD*=97) being faster than response to New Words (*M*= 553, *SD*=12).

As seen in the 2 X 2 ANOVA (Table 4.2) testing the effect of Reward (High v Low) and Word Type (Old v New) of confidence scores; there was also no difference in Confidence

---

[4] Note that fast RTs were expected during the recognition phase, since words were presented for 2200ms before responses could be given, and RTs were calculated from the start of the response screen. These fast RTs also indicate that participants were able to use the full colour scale within the allocated time window of 2700ms (it took 1659ms to move through the full colour scale from not confident to highly confident).

level for High or Low Reward Value words. Yet, there was a highly significant effect of Word Type on confidence. Confidence scores were higher for Old Words (*M*= 1426, *SD*=34) over New Words (*M*= 1124, *SD*=33; *t*(47)= 9.57, *p*<.001, *d*=0.90, BF_{10}= 4.51^{+8}). This finding suggests that participants were able to use the confidence scale appropriately and that the scale was sensitive to variations in confidence, but that reward did not influence subjective feelings of confidence.

4.4.2 Surprise Test*:* Subsequent Foil Recognition

Because final foil recognition was measured in the surprise recognition test with a Remember/Know (R/K) judgment on a six-point scale, I examined both overall recognition accuracy as well as the R/K memory experience for foils. First, I calculated overall recognition accuracy by collapsing across all correct 'Old' memory responses (Recollect, Sure and Unsure Familiar response given to Old items) and all correct 'New' responses (Unsure and Sure New responses given to New items). I then examined memory quality by calculating the proportion of responses to each category on the Remember-Know scale. Any 'Guess' responses on the Remember-Know scale were excluded from all analyses. Response Times were not of interest in the analysis of the surprise test as participants were instructed to take their time to accurately select the most relevant Remember-Know response, rather than focus on their response speed.

Table 4.3. *Surprise Test recognition scores for High and Low Value Foils alongside Discrimination scores (Pr).*

Reward level	Surprise Test Recognition	
	Old Words	Discrimination (*Pr*)
High Reward	.81 (.15)	.58 (.18)
Low Reward	.82 (.15)	.59 (.21)

Note: Values shown are Means (SD). *N*=48.

Recognition accuracy overall during the surprise test was good for both types of Foils (see Table 4.3) and New words (*M*= 0.77, *SD*=0.12). Overall there was a neutral Response Bias (*M Br*=0.57, SD=0.22; *Br* response bias was calculated for the average hit rates of all foil words using the FA rate for all New words).

4.4.2.1 Effect of Reward on Surprise Test Recognition

As seen in Table 4.3, subsequent recognition of foils did not vary due to Reward Value, as indicated by no difference in *Pr* discrimination ability (*t*(47)= 0.54, *p*=.590, *d*=0.05, BF_{10}= 1/5.853; *Pr* discrimination ability was calculated with one set False Alarms, therefore covaries in the same was as hit rates across within-subject conditions during the Surprise Test). Despite no effects of Reward Value on overall response accuracy, there could be more subtle effects of Reward Value on memory experiences, which would not be apparent when collapsing across all correct recognition responses. Therefore, the next analysis investigated whether Reward Value changed the profile of responses across the Remember/Know scale, by testing for interactions between Reward Value and the Remember/Know scale category for proportion responses. Any such interactions would suggest that the Reward Value manipulation had affected incidental encoding of foils in a way that resulted in differing subsequent memory experiences. There were however no such effects of Reward Value on Remember-Know scale use, since similar response patterns were given across the scale to both High Value and Low Value Foils. (*F*(2.94, 138.36)= 0.97 , *p*=.407*,* n_p^2*=.02;* For the Interaction term Mauchly's test of Sphericity was violated and a Greenhouse-Geiser adjustment was applies).

Self-reports

Post-experiment self-report questionnaires asked participants about effort and motivation levels during the study and initial test phases to investigate self-perceived differences in the High and Low Value Reward conditions. As well as examining any motivational or effort-based differences, it also provided a manipulation check to examine if the reward manipulation had an effect on motivation and effort levels. I expected to see increased motivation and effort for the High Value compared to Low Value Reward blocks. I also collected self-reported use of strategies, as I expected spontaneous strategy use to be more prevalent in the High Reward block, with the aim to increase reward gains disproportionally for High over Low reward items. A Likert scale of 1 to 10 was used where higher numbers indicated higher endorsement of that item, summary statistics can be seen in Table 4.4.

Table 4.4. *Self-report measures of motivation, effort and strategy use for the Study and Test phases in the High Reward and Low Reward Phases.*

Reward Level	Self-report Questions	Study Phase	Test Phase
High Reward	Motivation to gain reward	7.37 (1.63)	7.11 (1.81)
	Amount of effort	7.9 (1.47)	7.92 (1.30)
		Frequency	
	Strategy use	81%	-
Low Reward	Motivation to learn words	6.58 (1.83)	6.53 (1.78)
	Amount of effort	7.34 (1.51)	7.45 (1.55)
		Frequency	
	Strategy use	67%	-

Note. Motivation and effort values shown are Means (SD), and were collected on a scale of 1-10, with a higher score indicating more of the item measured. “Strategy use” refers to percent of participants who reported using a strategy to learn words in the study phases. *N*=38 for motivation and effort responses, N= 48 for strategy responses, since not all participants answered questions on motivation and effort.

Motivation and effort effects

As seen in Table 4.5, there were significant differences in Motivation and Effort levels between High and Low Reward conditions in the Study and Test 1 phases. Thus, despite no effects of reward on recognition memory performance, participants did report having higher motivation to gain rewards and exerting more effort during high than low reward tasks.

Table 4.5. *Paired-samples t-tests for self-report measures of motivation and effort for the Study and Test phases in the High Reward and Low Reward Phases.*

		High x Low Reward			
Self-report measure	Task Phase	*t*	*p*	*d*	BF_{10}
Motivation	Study	**3.69**	**.001**	**0.46**	**41.79**
	Test	**2.36**	**.024**	**0.37**	**1.99**
Effort	Study	**2.10**	**.043**	**0.32**	**1.23**
	Test	**2.13**	**.040**	**0.33**	**1.31**

Note. Significant values below .05 are shown in bold. *N*=38

Next, the difference between each participant's High and Low Reward recognition discrimination and bias scores on the first recognition phase were calculated and correlated with self-report measures of motivation and effort, to investigate if self-reports predicted changes in recognition performance at the individual level. Reward differences in self-reported motivation and effort on the first test were also correlated with surprise test recognition performance differences between high and low reward foils to investigate if individual differences during test 1 predicted reward effects of foil incidental encoding. As seen in Table 4.6, there were no relationships between effort and motivation levels and the difference in reward discrimination ability (*Pr*), response bias (*Br*) on test phase 1 or with reward discrimination ability (*Pr*) on the surprise test.

Table 4.6. *Spearman's correlations between the difference in self-report measures of motivation and effort between reward types with reward differences in recognition performance during test phase 1 and the surprise test.*

Self-Report Measures	Task Phase	Test 1 High-Low Reward *Pr*			Test 1 High-Low Reward *Br*			Surprise Test High-Low Reward *Pr*		
		r_s	*p*	$BF_{(10)}$	r_s	*p*	$BF_{(10)}$	r_s	*p*	$BF_{(10)}$
High-Low Reward Motivation difference	Study	.129	.37	1/2.77	.068	.686	1/4.13	.070	.676	1/3.98
	Test	.201	.226	1/2.16	.060	.721	1/3.75	.134	.423	1/1.86
High-Low Reward Effort difference	Study	.238	.15	1/2.64	.041	.808	1/4.65	.150	.368	1/2.51
	Test	.188	.258	1.57	.010	.951	1/4.43	-.063	.708	1/4.32

Note. All correlated variables are calculated based on a difference measure between High Reward and Low reward scores. (e.g. Motivation at study = Motivation during the High Reward study phase minus Motivation during the Low Reward study phase.). Bayesian correlations calculated with a Pearson's r width of 1). All correlations are based on *N*=38.

Since it has been suggested that strategy use is a method of enhancing performance to maximise rewards (Cohen et al., 2014; Middlebrooks & Castel, 2018), I also tested whether there were differences in strategy use between the High and Low Reward Study Phases (Table 4.4). Spontaneous strategy use was high for both conditions, yet participants were more likely to use a strategy in the High Rewarded (X^2(1)= 18.75, *phi*=.39, *p*<.001) than the Low Rewarded Study Phases (X^2 (1)= 5.33, *phi*=.11, *p*=.021).

4.4.2.2 Subjective Experiences of Recognition

The second line of investigation was related to item-specific processing and decision making. Within this experiment I was interested in how accuracy interacts with confidence and reaction times during initial retrieval attempts in terms of effects on incidental foil encoding, as measured by subsequent foil recognition. I therefore examined how trial-by trial differences in performance and confidence on Test 1 related to subsequent foil recognition on a within subjects' basis. Similarly to how I examined the effect of reward on subsequent recognition, I first investigated how these metacognitive and performance factors on Test 1 related to overall foil recognition by converting the Remember/Know responses into an overall accuracy rate (collapsing across all correct responses for 'Old and 'New' surprise test judgments). Then I followed up by investigating any differences in memory experience and quality by examining the distribution of responses across the Remember-Know scale categories as a function of Test 1 performance and metacognitive factors. For this analysis, all results were calculated collapsed across Reward conditions to increase trial numbers per condition, which is also parsimonious as no effects or interactions with Reward were seen in the earlier analysis. Only participants who had more than 10 trials per condition were included in this analysis, which resulted in data from 44 participants.

Surprise Test recognition task given Recognition Test 1 accuracy

The first question of interest was to see how accuracy, and related congruency of feedback affected subsequent recognition. I examined the proportion of correctly identified foil words in the surprise test, split by recognition phase 1 accuracy. Therefore, examining the proportion of surprise test foil hits, given that either a correct rejection (CR) or a false alarm (FA) was made about that foil word in the previous recognition test.

As a reminder, I expected that foils for which participants made errors (i.e. false alarms) on Test 1 would be recognised more often on the surprise test than foils that participants correctly rejected as new on Test 1, since the incongruent feedback may capture attention and elicit further processing of those items. This prediction was based on research showing that error-related feedback is more potent and surprising than congruent feedback (Fazio & Marsh, 2009; Metcalfe, 2017; Mullet & Marsh, 2016). When comparing overall surprise test hits, there was a highly significant effect of prior accuracy. Foil words that had previously received a 'false alarm' were more likely (supported by the Bayesian analysis suggesting extreme evidence in favour of the alternative hypothesis) to be later recognised as old on the surprise test than words that had received a 'correct rejection' on the first test (proportion accurate surprise test responses for all prior FAs; $M=0.87$, $SD=0.12$; all prior CRs; $M=0.79$, $SD=0.16$; $t(43)=4.82, p<.001, d= 0.55$; $BF_{10}= 695.29$).

Surprise recognition performance was not only calculated for overall accuracy but also split into subjective memory types (responses were given on a six-point scale); with memory for recognised words being judged as either Recollected, Sure Familiar or Unsure Familiar. Importantly with this analysis I was interested in exploring if the profile of responses differed depending on the Word Type: Prior FA or Prior CR, with expectations that there would be differences in memory experience related to if the word was previously correctly or incorrectly recognised. I therefore examined if the surprise test recognition memory experience for foils would differ depending on prior recognition test 1 accuracy, only focusing on the interaction of prior recognition accuracy (prior FA vs. prior CR) with the profile of responses spread across the Remember-Know scale. It would not be justifiable or statistically sound to examine any main effects here as these responses are not independent. There was a highly significant interaction between test 1 accuracy and the profile of the Remember-Know scale (F(4, 172)= 6.70, p<.001, n_p^2=.135). As is seen in Figure 4.4, where the profile of responses differs across the Remember-Know scale as a result of prior FA vs. prior CR.

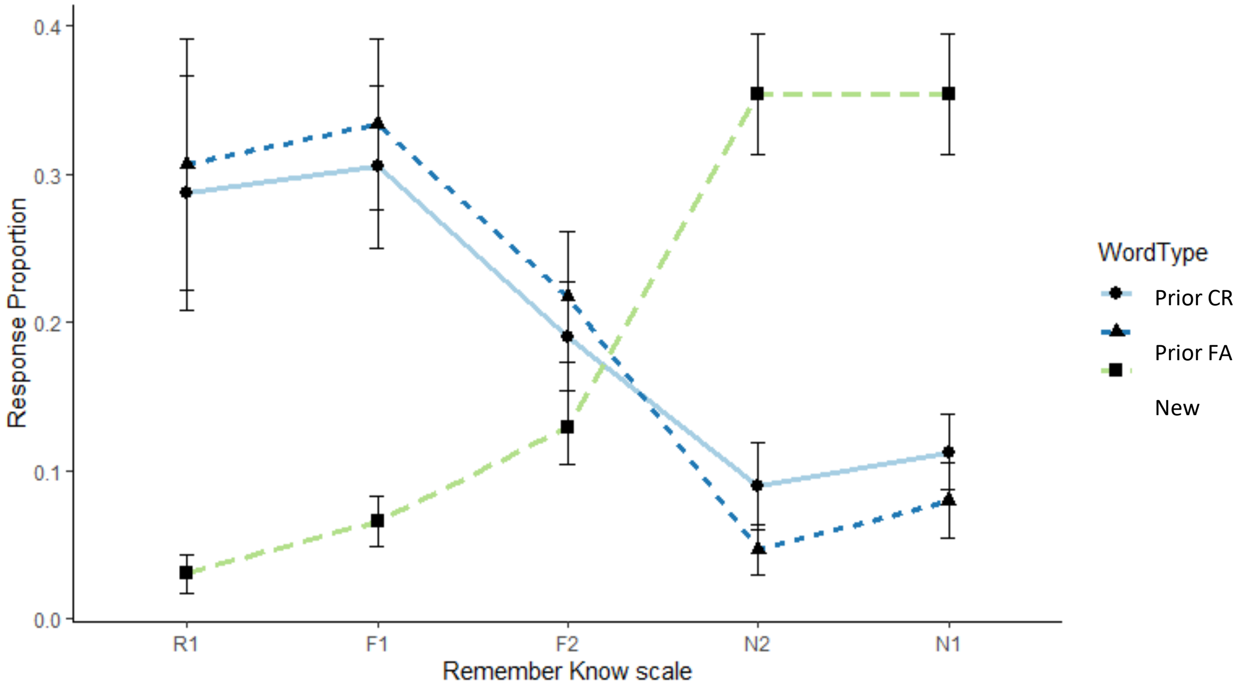


Figure 4.4. Proportion of surprise test responses on the Remember-Know Scale for Old foil words that previously received either False Alarms (FA) or Correct Rejections (CR) on Test 1, and for New words. The plot shows the mean proportion of responses for each word type and each scale response option along with the associated 95% confidence intervals. Note that Guesses (G) have been excluded in line with our analysis. The Remember-Know scale corresponds to: R1 = Recollect; F1 = Sure Familiar; F2=Unsure Familiar, N2=Unsure New; N1=Sure New. *N*=44

When examining the different pairwise proportions in scale responses, there was no significant difference in 'Recollect' responses, yet participants were more likely to respond 'Sure Familiar' for previous False Alarms compared to previous Correct Rejections as revealed in Table 4.7 (although note that this difference would not survive correcting for multiple comparisons). Furthermore, foils that previously elicited False Alarms elicited significantly fewer 'Unsure New' and 'Sure New' responses than foils that previously elicited Correct Rejections.

Table 4.7: *Pairwise comparisons of proportion responses between foils receiving prior False Alarms and Correct Rejections at each level of the Remember-Know scale.*

RK scale level		***t***	***p***	***d***	***BF_{10}***
Recollect	R	1.38	.179	0.07	1/2.23
Sure Familiar	F1	**2.43**	**.020**	**0.16**	**2.15**
Unsure Familiar	F2	2.00	.052	0.20	1/1.25
Unsure New	N2	**-3.19**	**.003**	**0.38**	**23.33**
Sure New	N1	**-4.21**	**<.001**	**0.54**	**22.39**

Note. Significant values below .05 are shown in bold. Guess responses were excluded prior to analysis. *N*=44.

The next analysis addressed a potential explanation for why errors in foil recognition on test 1 were recognised more often on the surprise test than prior correct foil identifications. Specifically, I addressed whether the increased recognition of foils that previously received false alarm responses was due to the incongruent feedback, which may have attracted attention and increased processing, resulting in enhanced incidental encoding of those foils (Butterfield & Metcalfe, 2006; Fazio & Marsh, 2009). As reviewed in the introduction, the effects of incongruent feedback on memory updating is heightened when feedback is unexpected, for example if people are confident about a memory decision and later get incongruent feedback informing them of an error, this mismatch is highly salient and leads to re-encoding of the correct information (Fazio & Marsh, 2009; Metcalfe, 2017). This explanation may be plausible for my results, since feedback was presented on the screen alongside the word for 1300ms after each memory decision, allowing sufficient time for additional encoding of foils that received surprising feedback.

One known moderator of feedback-related encoding is confidence. When incongruent feedback is given after confident responses, the feedback is more salient and leads to increased encoding. Although I was not able to test within False Alarms directly if high confidence led to increased subsequent memory compared to low confidence False Alarms as trial numbers were too low, I could however examine this issue within the Correct Rejections. I therefore investigated the effects of Test 1 Confidence, split by Test 1 Accuracy, on subsequent recognition. Importantly, feedback to low confidence Correct Rejections may also be surprising and salient as the congruent feedback would be less expected than congruent feedback for high confidence Correct Rejections. Therefore, if both foils receiving False Alarms on Test 1 and foils receiving low confidence Correct Rejections are recognised more often on the surprise test than foils receiving high confidence Correct Rejections, then this suggests that the surprising feedback may be driving the increase in subsequent recognition.

Subsequent recognition given T1 Confidence

Confidence on recognition Test 1 was significantly lower for False Alarms (M= 986ms, SD=292ms) compared to Correct Rejections (M=1113ms, SD=292ms; $t(43)=4.77$, $p<.001$, $d=0.43$; BF_{10}= 229.58), with extreme Bayesian evidence in support of the alternative hypothesis. Next, I split the Test 1 Correct Rejections into ‘High’ and ‘Low’ Confidence by calculating a median confidence value for each participant’s Correct Rejection trials. This value was then used to classify that participant's correct rejection trials into either High Confident (above their median) or Low Confident (below their median) on a within-subject basis. As seen in Figure 4.5, as would be expected this resulted in significantly lower confidence ratings for Low Confident (*M*=810, *SD*= 263) compared to High Confident Correct Rejections (*M*=1417, *SD*= 345; *t(*43)=20.41, *p*<.001, *d*=1.98; BF$_{10}$=2.61^{e+20}). Confidence was lower for False Alarms than for the High confidence Correct Rejections (*t(*43)=12.88, *p*<.001, *d*=1.35; BF$_{10}$= 9.13 $^{e+11}$) but higher for False Alarms than for Low Confident Correct Rejections (*t(*43)=6.47, *p*<.001, *d*=0.63; BF$_{10}$= 431.74). The differences in confidence between conditions was further supported by the Bayesian analysis showing extreme evidence towards to alternative hypothesis.

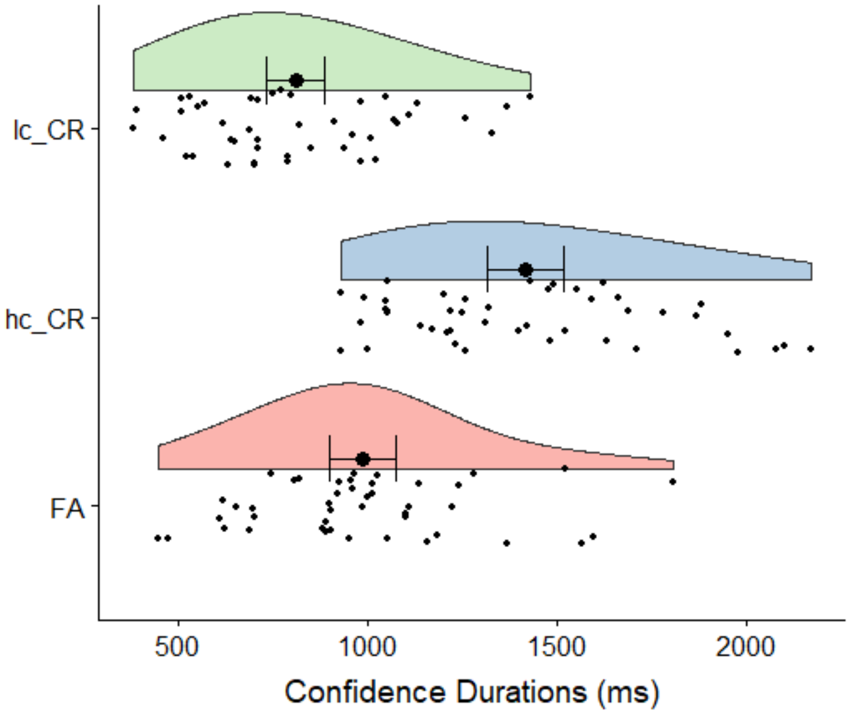


Figure 4.5. Distributions of Confidence measures for Recognition Test 1. The plot shows Confidence durations in seconds for Foil items, split by Accuracy (FA and CR) and for Correct Rejections, split again by High and Low Confident Decisions.

Subsequent recognition remained significantly higher for False Alarms compared to both Low Confident ($t(43)=4.41$, $p<.001$, $d=0.58$; BF_{10}, = 183.43) and High Confident Correct Rejections ($t(43)=4.61$, $p<.001$, $d=0.51$; BF_{10}, = 575.25), as seen in Table 4.8 and both supported by the Bayesian analysis showing extreme evidence towards to alternative hypothesis. There was however no significant difference in subsequent recognition between High and Low confident Correct Rejections ($t(43)=0.702$, $p=.486$, $d=0.09$; BF_{10},= 1/4.86). This finding suggests that surprising feedback was not the primary cause of increased subsequent recognition rates for foils receiving False Alarm responses on Test 1.

Table 4.8. *Proportion accuracy of subsequent recognition, split by Test 1 foil accuracy and for Correct Rejections also split by either high or low test 1 confidence, or slow or fast test 1 responses.*

Test 1 Foil Accuracy		Subsequent Recognition
Prior FA		.87 (.12)
Prior CR		.79 (.16)
Prior CR	Low Confident	.79 (.16)
	High Confident	.80 (.16)
	Slow Response	.78 (.16)
	Fast Response	.81 (.16)

Note. Scores shown means and SD in parenthesis. N=41

Subsequent recognition given T1 Response Time

An alternative reason for the increase subsequent recognition of False Alarms could be related to the increase initial Test 1 response times. The effect may be explained by increased initial processing of items when participants make slow responses, which in turn could increase incidental encoding. Again, trial numbers were too low to investigate the effects of slow and fast reaction times within False Alarm responses, therefore I investigated this issue for Fast and

Slow Correct Rejections on Test 1. I followed a similar procedure as the confidence analysis by sorting the trials within subjects, based on a median split for each participant's response times on Test 1.

Overall, Test 1 False Alarms (*M*=637ms, *SD*= 204ms) were slower than Test 1 Correct Rejections (*M*=559ms, *SD*=129ms; *t(*43)=3.34, *p*=.002, *d*=0.46; BF_{10}= 15.87). As intended, the median split produced Fast (*M*=318, *SD*= 88) and Slow (*M*=802, *SD*= 202) Correct Rejection conditions significantly differed in response time from each other (*t(*43)=19.05, *p*<.001, *d*=3.11; BF_{10}= 1.78 $^{e+18}$). False Alarms were associated with faster reaction times than the Slow Correct Rejection condition (*t(*43)=5.67, *p*<.001, *d*=0.81; BF_{10}= 3087.52) and slower reaction times than the Fast Correct Rejections (*t(*43)=13.00, *p*<.001, *d*=2.03; BF_{10}= 6.90 $^{e+15}$), both were supported by the Bayesian analysis showing extreme evidence towards to alternative hypothesis.. The differences in response times can be clearly seen in Figure 4.6. Importantly, if processing time is responsible for the increase in subsequent recognition for Test 1 False Alarms, then there should be similar increases in subsequent recognition rates for Slow Correct Rejections as seen with False Alarms, when compared to Fast Correct Rejections.

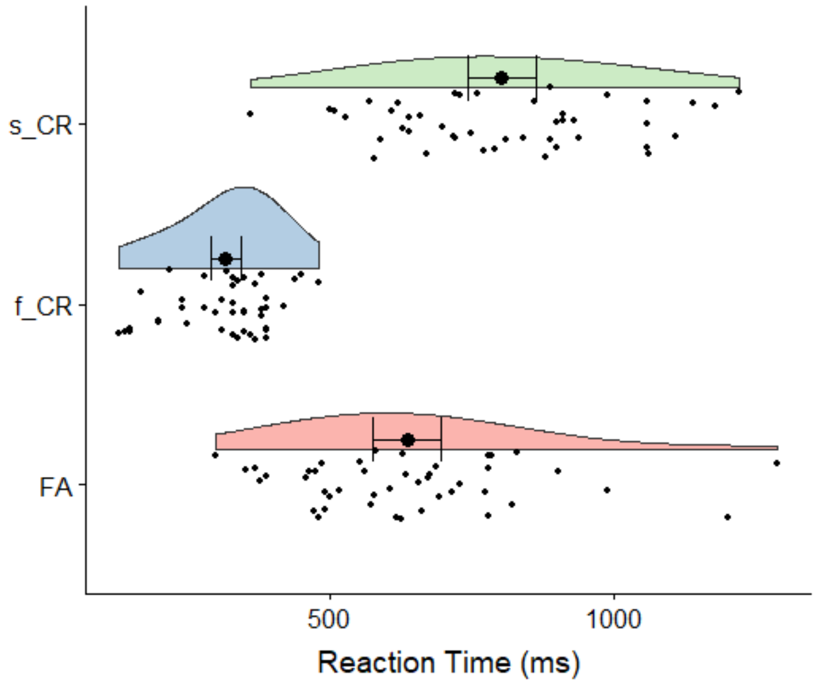


Figure 4.6. Distributions of Confidence measures for Recognition Test 1. The plot shows Confidence durations in seconds for Foil items, split by Accuracy (FA and CR) and for Correct Rejections, split again by fast (f.CR) and slow (s.CR) Correct Rejection Response Times.

As seen in Table 4.8, subsequent foil recognition remained significantly higher for Test 1 False Alarms compared to both Fast ($t(43)=3.43$, $p=.002$, $d=0.40$; $BF_{10}=15.73$) and Slow Correct Rejections ($t(43)=5.69$, $p<.001$, $d=0.68$; $BF_{10}=10038.16$). Interestingly, Slow Correct Rejections were *not* recognised significantly more than Fast Correct Rejections, because these conditions in fact showed a *reversed* pattern ($t(43)=3.29$, $p=.002$, $d=0.23$; BF_{10},= 15.81), contrary to a simple explanation that increased processing time enhanced encoding of foils.

In light of these additional results, it appears that feedback is not underlying the effect of increased subsequent recognition for Test 1 errors. There was no effect of congruency of feedback within correct foil identification on enhancement of subsequent memory. Similarly, longer processing times similarly did not explain the increase in subsequent recognition observed for False Alarms, since correct decisions made with longer response times did not enhance subsequent recognition. However, the additional analysis revealed that foils that received rapid Correct Rejections on Test 1 were recognised significantly more often on the subsequent foil test than Slow Correct Rejections.

4.5 Discussion

This study extended on findings from Experiment 1 where I found reward-related intentional effects on encoding, by examining the effects of external reward and recognition judgments on incidental encoding during retrieval attempts. This experiment also built on knowledge I gained in Experiment 2, where high value cues during recognition affected retrieval processing as indicated by reaction times and bias. Yet, in both these previous studies, I saw no effect of external reward on incidental encoding during retrieval attempts. Therefore, in Experiment 3, I attempted to enhance the effect of reward by adding metacognitive judgments and feedback, with the aim to increase the salience of rewards. As in the previous two chapters, I similarly used a memory-for-foils paradigm (Jacoby, Shimizu, Daniels, et al., 2005) which enabled me to assess if such reward-related enhancements of memory interacted with meta-memory processes during retrieval attempts, as well as enabled me to investigate metacognitive factors separately from reward. Thus, this experimental design allowed me to examine if different levels of external reward and metacognitive processes that are present during the first test, modulate incidental encoding of the novel foil words. This is similarly measured by subsequent recognition on the surprise test.

Overall, I found that external rewards presented during the encoding and retrieval phase had no effect on intentional encoding or subsequent incidental encoding, with no differences due to the high and low value reward conditions on any performance measures. This is contrary to what I saw in Experiment 1, where there was a modulation of intentional encoding due to external rewards. The lack of intentional modulation was surprising as the experimental design was similar to that in Experiment 1. On reflection, in this experiment, there were much longer word lists. As word list length increase there is more risk of interference effects and increased errors (Underwood, 1978). There is a slight decrease in performance between here as compared to Experiment 1. However, I was not expecting this to remove the effect of rewards on recognition. Specifically, in this experiment, where motivation manipulation was critical, the increased word list lengths could have had an effect on participants' ability to sustain motivation and attention, which the reward enhancement relies upon.

Another important difference to Experiment 1 was that type of reward given was changed from monetary to food-based in the current study. This change was done for two reasons. Firstly, to create a more automatic response to the rewards, and secondly to reduce the financial costs associated with providing monetary rewards. Food rewards (primary) have been seen to be equally as rewarding as monetary rewards (secondary), however there is evidence suggesting that these rewards are processed in different ways (Beck et al., 2010). In Beck *et al*'s study however, the primary reward was administered trial-by-trial and reinforced behaviour, whereas in this experiment, food rewards were accumulated. The chocolate rewards would then have been treated in a more abstract way, similar to monetary coins, but inherently would hold less value than actual financial gains. Overall, the chocolate reward manipulation in this experiment did not boost intentional, let alone incidental encoding processes.

Even though participants did report that they were motivated by the high reward condition, the failure in manipulating memory as a function of reward may be due to how reward was presented during the procedure. Reward cues were presented at the beginning of each study and test blocks, with no trial-by-trial reward cue, which could have reduced the saliency and importance of the reward within each trial. This was similarly the case in Experiment 1 however, the test phases were substantially shorter than in this experiment, where the addition of confidence ratings and feedback substantially increased the inter-trial interval.

The link of the reward within each trial could have lost salience and importance as the test trials progressed.

Linking to the lack of trial-by-trial reward cues is an important limitation, which was the lack of reward indicative feedback. Instead feedback indicated accuracy, and so was not explicitly reinforcing the rewards received explicitly and may have taken the participant focus away from the rewards and onto the trial accuracy. An improvement to the current design would be to explicitly link the feedback to reward winnings, providing trial-by-trial reward earning. This would enhance the salience of the reward manipulation, in line with prior research that indicated ongoing winnings (Adcock et al., 2006; Beck et al., 2010; Murty & Adcock, 2014). When examining feedback more generally, we cannot determine if feedback was boosting encoding compared to if no feedback was given. As in this paradigm there was no manipulation of feedback, with feedback being given on every trial. This is therefore impossible to assess if feedback itself had an effect or if there was an interaction of feedback with value. Feedback generally is suggested to increase attention and salience of information (Pashler et al., 2005). Feedback should be more potent in highly rewarding, compared to no or little reward conditions. This is something that could be examined with the view to increase the potency of rewarding information.

Even though effects of rewards on the memory for foils were absent, interestingly, a very strong effect was seen when examining subjective judgments of recognition. In particular, accuracy on the first test predicted subsequent foil accuracy on the surprise test. With foils that elicited false alarm errors (responding “old” to new items) on test 1 being recognised more accurately on the surprise test, compared to foils that elicited correct rejections (responding “new” to new items) on test 1. This finding was produced by more “sure familiar” responses and fewer “sure and not sure unfamiliar” responses on the RK scale during the surprise test for prior false alarmed foil words than for prior correctly rejected foil words. Interestingly however there were no significant differences between these conditions in proportion of "recollect" responses.

The increase in "sure familiar" responses suggest that there was a stronger memory for items that had previously elicited a false alarm error on test 1. From this paradigm I cannot determine if these effects were due to a real false memory of previously studying those items or not. But, the lack of difference in "recollect" responses here suggests that there was not a qualitative difference in the 'type' of the memory between prior false alarms compared to prior correct rejections. This would be something that I expected if there was a distinct experience of studying those false alarmed items as recollection is defined by being able to remember contextual details or aspects of the memory episode (Donaldson, 1996).Whereas in this study both prior false alarms and prior correct rejections elicited a feeling of familiarity to differing strengths. Potentially there was extra-experimental interference or a sustained false familiarity, resulting in a sure experience in previously seeing the falsely alarmed items, but this was more a feeling of knowing, rather than a clear experience of studying those items. This was supported by a decreased proportion of incorrect subsequent recognition responses for false alarms. Again, this supports the view that there was a overall sense of familiarity for the false-alarm errors.

Unfortunately, in this study I can only speculate about what this means for the underlying processes driving these differences. Currently there could be an underlying extra-experimental familiarity for false-alarm items, or there could indeed be a false memory for studying those items, not indicated by the 'Remember/Know' paradigm. Just as importantly it cannot be seen here if there is a boost in encoding for items that were false alarmed to or if the increase in subsequent recognition was a sustained effect. Participants felt that they had already seen these items before so this could have been carried over to the surprise test, or there could have been a boost in falsely ‘re-encoding’ these items after the first recognition judgment had happened.

The current paradigm looks at both accuracy as well as confidence and response times. In order to determine what processes might be underlying this significant effect of initial judgement accuracy on later foil recognition, I considered that there were significant differences in confidence ratings between these conditions on the first test. Correct rejections were associated with higher confidence overall than false alarm errors. Thus, confidence may have affected foil accuracy encoding either in a simple way (lower confidence = enhanced encoding due to e.g. enhanced attention and monitoring), or by interacting with the feedback to enhance encoding particularly when feedback was surprising (low confident correct responses receiving congruent feedback, or high confident but erroneous responses receiving incongruent feedback (Butterfield & Metcalfe, 2006; Fazio & Marsh, 2009; Metcalfe, 2017)). However, neither of these accounts were supported, because when correct rejections were split by high and low confidence there were no differences between the high and low confident correct rejections in terms of subsequent foil recognition. This result thus suggests that confidence may not be underlying the effect whereby foils associated with false alarm errors on the first test were recognised better on the later surprise tests compared to correctly identified foil words on test 1. Importantly to note, the hypercorrection studies examining feedback and confidence have primarily looked at semantic memory (Butterfield & Metcalfe, 2006; Fazio & Marsh, 2009; Metcalfe, 2017), where there is a factual correct response. When transferring this to episodic memory literature I was similarly expecting to see a similar effect. Yet, in the current study, due to limited trial numbers, the hypercorrection effect could not be investigated in the same way as prior studies suggested. False alarm errors were not primed and therefore errors varied between participants, with some participants attaining very few false alarm errors. Therefore, false alarm errors could not be split into reasonably sized high and low confident subgroups to assess the difference between confident and non-confident errors. Looking at the effect of confidence on encoding of new items provided some insight with no increase in encoding of low compared to high confident correct rejections as discussed above.

Another underlying factor could be response times during the initial recognition test, where participants were asked to respond as quickly as they could after viewing the word for 2200ms. Response times to make this ‘Old/New’ recognition decisions during the initial tests varied significantly depending on accuracy once again, with participants taking significantly longer to make false alarm errors than correctly rejecting foils. Therefore, it is reasonable to expect that the enhanced subsequent recognition of false alarms could be due to participants having more time to engage in elaborate processing of words eliciting false alarms than correct rejections during the initial recognition tests. Elaborate processing is more effortful and takes longer to engage with, compared to more automatic decisions (Yonelinas, 2002). However, when I examined this by splitting correct rejections based on reaction time and comparing subsequent foil recognition, I found that in fact foils associated with faster correct rejections on test 1 were later recognised significantly more accurately than foils associated with slower correct rejections on test 1. Hence, the subsequent differences between false alarms and correct rejections on the final test cannot be due to a simple relationship with processing time on recognition test 1.

Interestingly therefore, the results suggest that there may be two different processes occurring during foil processing on an initial recognition test that enhances later recognition memory for foils. The first process may be related to the enhanced detection of the novelty of some foils, which leads to quick recognition judgments and better encoding than for foils that are not as quickly classified as new. The second process seems related to errors in detecting which words are previously encountered, in that foils that are misjudged as old take longer to be classified on the first test, but are recognised on the final foil test more often. Therefore, it may be the case that foil words that elicited false alarm errors on the initial recognition test could be associated with some long-term familiarity or extra-experimental source of memory signal (Bridger, Bader, & Mecklinger, 2014). Such familiarity seems to impair decision accuracy on the first test when those items are "new" in the experimental context, but facilitates decision accuracy on the surprise test when those items are "old" in that experimental context. It is therefore relevant to examine how long-term familiarity or extra-experimental influences interact with episodic encoding and retrieval processes to determine how a stimulus is judged as 'old' or 'new' and how this affects learning during this retrieval decision.

Familiarity and novelty detection can be argued to be 'two sides of the same coin'; where there is a lack of familiarity there is novelty, and a lack of novelty is due to an increased sense of familiarity (Rugg & Curran, 2007). This interpretation is in line with signal detection theory (Wixted, 2007), where recognition decisions are determined by memory strength exceeding a criterion that determines if a stimulus is 'old', otherwise the stimuli is judged as 'new'. Similarly, the pattern of RTs and Confidence fits with signal detection theory, with memories that are further from the criterion in either direction (more familiar or more novel) being more confident and more rapidly recognised.

Other lines of research have suggested potential dissociable processes for novelty and familiarity detection (Daselaar, Fleck, & Cabeza, 2006). Novelty detection is argued to be driven by activation in the anterior hippocampus and rhinal cortex, and familiarity is driven by activation of the posterior parahippocampal gyrus (Daselaar, Fleck, & Cabeza, 2006). Therefore, familiarity and novelty signals may be interacting to determine recognition judgments, either being present or absent in some cases or in fact be two complementary processes. Before I can confidently attribute the dual enhancement in subsequent recognition differences for foils to prior error making and novelty detection however, I need to explore more in depth what processes during recognition test 1 are occurring that are leading to these marked differences.

Notably this experiment has raised interesting questions about how ‘Old/New’ recognition decisions are made, and in turn how these decisions and processes lead to incidental encoding differences for new information encountered during recognition attempts. However, from this study alone I cannot fully address these questions. Two main points of interest arose related to memory for foils; firstly, that foils that received false alarm errors were recognised most often on a subsequent recognition task, and that foils associated with fast correct rejections were also subsequently recognised more often than other slower correct rejections. The questions therefore then surrounding these findings relates to what are the underlying neurocognitive processes underlying these effects, as investigated in the subsequent experiments.

4.6 Conclusions

Overall, this chapter adds to the existing knowledge of modulating factors that influence incidental encoding during retrieval attempts. Spontaneous incidental encoding occurred at a high rate even without a processing manipulation during the intentional encoding stage, which classically have been applied in the memory-for-foils literature, for example using a levels-of-processing manipulation (Danckert et al., 2011; Jacoby, Shimizu, Daniels, et al., 2005). Overall, the strongest predictor of foil recognition on the surprise test was false alarms on the first recognition test; responding “old” to new ‘foil’ words. This finding suggests that there was a failure to detect novelty in these items, and instead these items elicited a persistent familiarity memory response that carried over to the final test. When looking deeper into the effects of accuracy and response characterises it becomes apparent that there is not a clear relationship between response characterises such as response times and confidence judgments, on incidental encoding. Rapid responses, similarly to incorrect responses, showed increased subsequent recognition rates. This indicates multiple processes occurring during recognition judgments that result in different levels of incidental encoding. Therefore, overall, the findings demonstrate the importance of extra-experimental factors that affect incidental learning such as false familiarity (false alarm errors were recognised more on the surprise test) as well as other recognition judgments (rapidly responded to correct rejections were recognised more on the surprise test). This study starts to provide more understanding about the effects of recognition judgments and specific underlying processes, such as the interplay between novelty and familiarity detection.

It is known from prior literature that memory monitoring plays a role in recognition judgments (Finley et al., 2010). Relating to this, the current study provided knowledge about how metacognitive factors interact to affect incidental encoding. Somewhat surprisingly, self-reported measures of confidence did not appear to relate to subsequent recognition rates. Yet, analysis of reaction times provided tentative evidence that simple processing time was not responsible for the increased subsequent recognition of prior false alarm errors, and provided insight into a complementary process of novelty detection. Overall, additional research is needed to examine if familiarity and novelty detection are two separate processes that are engaged during recognition attempts, and how they relate to differences in incidental encoding of new information.

Moving forward in light of the results from this chapter, an adapted paradigm will be used to investigate neural dissociations between familiarity and novelty processes with EEG. Previous EEG findings have found reliable event-related potential (ERP) effects for familiarity processes and novelty detection (Rugg & Curran, 2007; Friedman et al., 2001). Therefore, it would be interesting to investigate familiarity and novelty processing with EEG in the current paradigm. This would allow me to separate the neural processes driving false alarm errors from correct rejections, and also explore effects of fast versus slow correct rejections as well as high and low confident correct rejections. The examination of ERP activity will enable me to determine if the relationship between foil recognition and accuracy and metacognitive processes that can be observed behaviourally are driven by qualitatively different neural processes, which will add knowledge to the debate to whether recognition memory involves dual or recently suggested triple dissociable processes (Daselaar, Fleck, & Cabeza, 2006), or alternatively whether all foil recognition effects can be explained by a single memory process, that gives rise to a continuum from familiarity strength to novelty detection.

4.7 Summary

Experiment 3 provides knowledge about what processes are contributing to retrieval attempts during recognition judgments and what processes might be affecting incidental encoding rates. Multiple processes occurring during recognition indicate the complexity of recognition tasks, with the interaction of false familiarity and novelty being highlighted here. These findings converge with prior arguments about dual and triple dissociations of recognition processes (Daselaar, Fleck, & Cabeza, 2006; Yonelinas et al., 2010). Yet, as this experiment only provides behavioural measures it is impossible to speculate too much about what neural processes are driving the effect which leads to increased false alarm recognition as well as correct rejections that were rapidly responded too. This study starts to provide knowledge about these processes and how these both contribute to recognition judgments. In the next chapter I will go on to investigate the neural processes underlying false familiarity and novelty detection with ERPs, and examined the role of these processes in incidental encoding during retrieval attempts. These effects will be both examined in a younger population and in a healthy older adult population. With this ageing factor, I aim to examine if older adults perform differently to younger adults in relation to the influence of false-familiarity and novelty on incidental encoding, and how the underlying neural recognition processes may differ across ages.

Chapter 5: Electrical brain activity associated with false memory and novelty-related increases to subsequent recognition. Exploring the effect of healthy ageing on recognition processes.

5.1 Overview

In this chapter, I extended on Experiment 3 in Chapter 4 where I hypothesised about what neural processes are involved during recognition decisions for new items in a recognition test that lead to differences in subsequent foil recognition. I suggested that there were two distinct neural processes; familiarity and novelty detection that interact during recognition decisions, and that these affect incidental encoding of foils in different ways. However, it is not possible to test these hypotheses with behavioural methods alone as in Experiment 3. Therefore, in the next experiment I modified and improved the paradigm in Experiment 3 to test for hypothesised effects with ERPs. In this ERP experiment I also extended on my previous research questions to investigate these predicted effects in both younger and older adults. This study therefore followed up on the findings from Experiment 1 (Chapter 2) also, where older adults matched younger adults in their ability to constrain memory search and subsequently recognise foil words. Even though the older adult group surprisingly did not show memory deficits in Experiment 1, I could not make firm conclusions about if this group had fully comparable memory processes to the younger group, or alternatively, if this group relied on compensatory mechanisms or different retrieval processes to solve the task. Therefore, in this chapter, I also compared between-group differences in ERP activity related to recognition processes after intentional and incidental encoding, as well as tested my hypothesised processes underlying foil encoding in healthy ageing.

5.2 Experiment 4

In this chapter, I investigated the item-specific neural processes that lead to differences in incidental encoding of new items, and potential neural differences across age groups. By using the memory-for-foils paradigm (Jacoby, Shimizu, Daniels, et al., 2005) with ERPs, this provided a suitable method to separate different retrieval processes that contribute to recognition decisions, to measure item and age differences in neural activity between old items being recognised, new items being correctly rejected, and also new items being falsely recognised on the initial recognition test. Then, I looked at the after-effects of different recognition and decision processes during the initial encounter of foils on subsequent recognition-related ERP activity for those foils that had been incidentally learned. Both these investigations were conducted with younger and older adults, to investigate potential age differences in retrieval processes and encoding of foils.

The motivation for this chapter came from Experiment 3 in Chapter 4, which exposed unexpected insights into recognition processes that had an effect on subsequent recognition rates. In Experiment 3, responses to foil items during the first recognition test seemed to induce incidental encoding at different rates. Foils that elicited false alarm errors (responding ‘old’ to new items) on Test 1 were recognised more accurately on the surprise subsequent recognition task, as well as foils that were associated with faster correct rejections on Test 1. However, firm conclusions could not be drawn about why foils associated with false alarm errors were recognised more often on the surprise test than foils eliciting correct rejections, nor why fast correct rejections appeared to boost incidental encoding of those foils compared to slow correct rejections. I made suggestions about potential processes driving these effects, but in this chapter, I directly tested these hypotheses.

In Experiment 3, increases in subsequent recognition of foils that received false alarm responses on Test 1 could have been due to sustained false memory for those items, perhaps caused by an extra-experimental source of familiarity resulting in repeated familiarity experiences for those items, or alternatively due to increased incidental encoding of foils that received incongruent error feedback (Butler et al., 2011; Butterfield & Metcalfe, 2006; Metcalfe & Finn, 2011; Mullet & Marsh, 2016). The use of continuous, accurate feedback however prevented me from assessing the role of feedback in foil encoding in Experiment 3, as feedback congruency would have needed to be manipulated separately from response accuracy in order to test a causal relationship. Unfortunately, the additional 'Remember/Know' responses in Experiment 3 also did not shed much light on the underlying processes giving rise to differences in foil memory. The main differences seen between conditions were in the 'sure familiarity' rather than 'recollection' and 'new' response categories.

Therefore, a better way to examine the underlying cognitive processes driving memory for foils is to directly assess where an item is remembered from by collecting source memory judgments (Bröder & Meiser, 2007; Johnson et al., 1993; Mitchell & Johnson, 2000). Tests of source memory provide more detailed information about the nature of memories, such as what contextual details people can remember (Bröder & Meiser, 2007). Source judgments in the foil recognition test could provide information about if subsequent recognition of foils receiving false alarm responses on Test 1 is driven by a false memory of those words from the original study phase, or alternatively, a correct memory of seeing foils only from the first test phase. Therefore, source judgments, coupled with removal of feedback in Experiment 4 enabled me to assess more directly if foils associated with Test 1 false alarms were recognised more often on the final test due to a persistent false memory, or due to enhanced encoding during the first test.

The second interesting result in Experiment 3 that I will be examining here was the effect of correct rejection response times on subsequent recognition. In Experiment 3, fast correct rejections for foils on Test 1 resulted in increased correct subsequent recognition compared to slower correct rejections for foils on Test 1. A similar analysis was conducted on high vs. low confidence ratings, but no effect of Test 1 confidence on foil encoding was seen in Experiment 3. Reaction times and confidence ratings can be useful in providing insights into the retrieval and decision processes underlying recognition judgments (Ratcliff & Starns, 2013; Weidemann & Kahana, 2016). However, a confound in the Experiment 3 paradigm limited the interpretability of this effect, as response time and the confidence scale were time-limited, which meant that very slow responses could never hold high confidence. Firstly, as false alarms were slower on average, it was difficult to assess if all false alarms were actually lower in confidence than correct rejections overall or if due to slower response times, the full confidence scale was not usable. It was also not possible to accurately assess if fast and slow correct rejections differenced in confidence, as slow correct rejections confidence responses were bound by the response time window. Therefore, in this chapter the confound of having confidence bound by response times has been addressed after piloting changes to the paradigm (See Appendix 4.1.2). The improved Experiment 4 paradigm could therefore be used to assess the effects of Fast and Slow, as well as High and Low Confident Correct Rejections[5] and the associated neural activity that contributed to these types of decisions.

The main aim of Experiment 4 was to examine what neural processes occur during an initial old/new recognition test (Test 1) when foil items are being initially processed, and how those processes relate to subsequent recognition of foils as assessed with a subsequent old/new recognition test for foils (in a surprise test). I expected to replicate the behavioural effects I saw

---

[5] Although note that confidence did not influence behavioural memory performance or neural activity for foils, and therefore the analysis of high vs. low confidence correct rejections is not presented in this chapter.

in Chapter 4, with increases in subsequent recognition for foils associated with prior False Alarms, as well as for prior Fast Correct Rejections compared to prior Slow Correct Rejections. In light of the results from Experiment 3, I predicted that two processes could be occurring during the recognition of foils that both lead to an enhancement of subsequent recognition: false familiarity and novelty detection, which would converge with the wider literature on recognition processing (Curran, 2000; Friedman et al., 2001; Paller et al., 2007; Rugg & Curran, 2007; Wilding & Rugg, 1996). Therefore, I aimed to investigate if different neural correlates could be seen in the ERP data for false familiarity and novelty detection on Test 1, and the role of these neural processes in relation to incidental encoding of foil items. Additionally, I examined these effects in both younger (18-28) and older (>60) age groups to explore if these neural processes are similar or different in a healthy ageing population. I expected the two age groups to have a similar behavioural memory performance, given that in Experiment 1 (Chapter 2) I found almost identical memory ability between similar participant age groups in a very similar task.

The neural processes of relevance in this study are familiarity, recollection and novelty detection, all of which are known to be involved in deciding if an item is ‘old’ or ‘new’ during a recognition task (Curran, 2000; Daselaar, Fleck, & Cabeza, 2006; Friedman et al., 2001; Paller et al., 2007; Yonelinas, 2002). As previously reviewed in Chapter 1, familiarity is associated with a ‘feeling of knowing’ that a stimulus has been encountered before, and can vary in strength (Paller et al., 2007; Yonelinas, 2002). The ERP correlate of familiarity, the FN400 or mid-frontal old/new effect, occurs around 300-500ms after a familiar stimulus is presented (Curran, 2000; Curran & Hancock, 2007; Rugg & Curran, 2007; Yonelinas, 2002) with increased familiarity experiences associated with a reduced mid-frontal negativity (i.e. more positive ERPs for familiar than unfamiliar stimuli, Curran & Cleary, 2003). I therefore expected to see increased familiarity-related ERP activity for False Alarm errors compared to all types of Correct Rejections, with a reduced negativity in the early time-window around the mid-frontal regions in the younger adults.

Familiarity processing is arguably spared in healthy ageing populations, with a range of behavioural and neural evidence to support this view (Daselaar, Fleck, Dobbins, Madden, & Cabeza, 2006; Fraundorf et al., 2019; Friedman, 2013; Koen & Yonelinas, 2016; Yonelinas, 2002). However, it is worth noting that evidence about the role of the FN400 component in recognition memory for older adults is mixed, with some findings showing diminished (Duarte et al., 2006; Guillaume et al., 2009; Wolk et al., 2009) or even absent (Wang et al., 2012) FN400 effects, especially for word stimuli (Ally et al., 2008), suggesting that the FN400 may not indicate familiarity in older adults (Wolk et al., 2009). I am therefore cautious to suggest any directional effects in FN400 between the age groups given the mixed findings in the current literature set.

Recollection is argued to be a qualitatively different memory processes to familiarity, and is the ability to recall distinct contextual details about a previous encounter with a stimulus (Paller et al., 2007; Rugg & Curran, 2007). The ERP correlate of recollection, the left parietal old/new effect, is maximal across the left parietal scalp around 500-800ms after stimuli onset (Rugg & Curran, 2007). The left parietal old/new effect has been reliably observed in younger adults, with increased positivity for items which are correctly recognised (Hits) compared to items that are identified as new, especially when recognition is associated with a feeling of recollection or an ability to remember source/contextual details about items. I therefore expected increased parietal positivity in this time-window for hits compared to correct rejections, both after intentional (at Test 1) and after incidental encoding (at the Surprise Test) in the younger adults.

The late parietal old/new effect however has not always been observed in healthy older adults (Friedman, 2013; Guillaume et al., 2009; Horne et al., 2020; Wolk et al., 2009) with the changes or lack of this neural activity supporting the reduction in recollective memory experiences in older age, which converges with the behavioural literature (Koen & Yonelinas, 2016; Yonelinas, 2002). Some studies however have found that there is intact ERP evidence of recollection in older adults if this effect is related to high performance (de Chastelaine et al., 2016; Duarte et al., 2006; Wang et al., 2016) whereas others show a reduced or diminished effect (Wang et al., 2012), or intact behavioural recollection but no ERP effect (Horne et al., 2020; Nessler et al., 2008). From these mixed findings, it is apparent that there is not a definite absence of recollection in older age, but that other factors may play a role in determining whether the ERP correlate of recollection is observed in older adults, such as memory performance and paradigm requirements. However, since generally recollection is thought to be impaired in older adults, I expected to see differences in this later time window when compared to younger adults, with the expectation that there will be a diminished late parietal old/new effect in the older adult group after intentional encoding (Friedman, 2013; Guillaume et al., 2009; Horne et al., 2020; Wolk et al., 2009). I am however more cautious to suggest directional effects of ageing on the parietal old/new effect after incidental encoding, given the limited literature to draw on using incidental encoding (Wagnon et al., 2019), but as incidental encoding here is reliant on initial recognition judgments, recollective based activity here may be more linked to test 1 recognition success.

The P3 effect occurs at a similar time to the parietal old/new effect and has a similar parietal topography, indicating a potential overlap in processes around the 500-800ms time window (Barry et al., 2020; Friedman et al., 2001; Polich, 2007; Ranganath & Rainer, 2003). The P3 component may index orientation of attention towards novel items, and reduces in magnitude due to repeated exposure indicating a habituation of the stimuli's novelty (Friedman et al., 2001; Knight, 1996). Therefore, I would expect a heightened P3, with increased parietal positivity for fast correct rejections if this response is being driven by increased novelty detection for those items. Such an orienting effect may explain the subsequent increase in recognition for fast over slow correct rejections (Cycowicz & Friedman, 2007) as similar differences in ERP positivity have been seen to predict successful encoding and thus subsequent recognition (Friedman & Trott, 2000). With the reasoning from the subsequent memory effect seen in the 'DM' literature, showing that differences during recognition, for example between 'remember' and 'know' items or between the processing of words and non-words, should be evidenced during encoding (Friedman & Johnson, 2000; Friedman & Trott, 2000; Otten & Rugg, 2001; Otten, Sveen, & Quayle, 2007).

Interestingly, there are numerous studies suggesting that the magnitude of the novelty P3 is intact in healthy older adults (Berti et al., 2017; Daffner et al., 2005; Friedman et al., 1998), with age differences arising due to differences in novelty habituation rather than initial novelty responses to a first encounter with new items (Richardson et al., 2011). Therefore, I would expect to see similar P3 enhancement for fast compared to slow correct rejections in the healthy older adult group as in the younger group.

To my knowledge, these item-specific retrieval processes have not yet been explored in terms of how they relate to incidental encoding of novel information during retrieval attempts. The current research therefore aimed to examine the dissociation between novelty and familiarity processes and provide a suitable paradigm for investigating the neural basis of these processes with EEG. Since my version of the paradigm involved recognition testing after both intentional and incidental encoding, it also provided an excellent opportunity to examine how healthy ageing affects these different types of memory processes, which is important given the limited research into incidental encoding in older adults (Wagnon et al., 2019).

5.3 Methods

The Experiment 4 paradigm was similar to that used in Experiment 3, but adapted to resolve the previously mentioned potential confounds (confidence bound by reaction times). In Experiment 3, the response window was capped per trial to 2700ms, meaning that responses after 1665ms would have never been able to elicit a maximally high confidence response. In Experiment 4, I removed the fixed response window to enable a separation of confidence and Reaction Times. Further changes were to remove the trial-by-trial feedback during the initial recognition test, to reduce the potential processes that could have caused subsequent increases to false alarm errors. I also removed the reward manipulation that was used in Experiment 3 (since that manipulation did not affect behaviour), and replaced this with a constant reward throughout the task to ensure task engagement. The final change was to replace the Remember/Know judgments for foils on the final surprise test with source judgements, to provide more insight into the processes behind the increased final recognition of foils that received prior false alarm responses on Test 1. A pilot study validated these paradigm changes and confirmed the new version would be suitable for use with ERPs (see Appendix 4.1.2).

Experimental methods, data analysis, participant recruitment and exclusion followed pre-registered criteria. All criteria were submitted to the Open Science Framework (osf.io/w4x6h) along with the key behavioural and ERP predictions.

5.3.1 Participants

I aimed to recruit 36 participants per group in order to achieve >0.9 power to detect an effect size of Cohen's d=0.7 at two-tailed α=.05, and to fully rotate and cross all counterbalancing factors (task orders, stimuli assignment to conditions, etc.). The final sample was reduced to 32 for Younger Adults (*M* age = 21.41years, *SD*=2.40, range = 18-28; *M: F*= 13:19) and 30 for Older Adults (*M* age = 69.43years, *SD*=6.99, range= 60-89; *M: F*= 11:19). This sample size was fairly similar to prior EEG studies (Vogelsang et al., 2018), with the smaller final sample partly due to practical constraints and participant exclusions, resulting in >0.99 power to detect a large effect size of Cohen's d=0.8 at two-tailed α=.05 or a >0.97 power to detect a medium effect size of Cohen's d=0.5 at two-tailed α=.05. Specifically, an additional nine Younger Adults were replaced or excluded (seven due to low False Alarm trial numbers and three due to poor EEG quality after pre-processing). An additional 16 older adults were also excluded (nine due to low False Alarm trial numbers and six due to poor EEG quality after pre-processing). The Younger Adults were recruited from the University of Kent and took part in the study in return for course credit (*N*=4 of final sample) or £20 reimbursement (*N*=28 of final sample) for their time. The Older Adults were recruited through local advertisements and through the Kent Psychology Community Mailing List. The Older Adults all received £20 reimbursement.

All participants were native English speakers, right-handed, had normal or corrected to normal vision (colour-blindness was not tested and only self-reported), reported not having depression, anxiety or dyslexia, and were not taking any psychoactive medication. All participants were reimbursed for their time and 'won' two small treat size bags of milk chocolate or a dark chocolate or lactose-free alternative. This was not advertised but used in the task to maintain task engagement. Participants had not taken part in any other experiments within this thesis. All participants provided full written consent and the procedures were approved by the University of Kent School of Psychology Ethics Board. Participants were mostly tested by LS with some Older Adult participants being tested by two MSc students who were trained by LS and all testing was conducted in a face-to-face manner.

The Older Adults had scores within the normal range on the Montreal Cognitive Assessment (Nasreddine et al., 2005; MoCA; cut off >= 25[6]; Older Adults *M* =28, *SD*= 1.55) and had a high level of vocabulary understanding (Shipley Vocabulary: *M*=35.70 , *SD*=1.73, Range = 30-37). Older adults were highly educated (participants' education level; GCSE=2, A Level/equivalent=4, Undergraduate=8, Postgraduate=9, Other qualifications=5). MoCA, Shipley Vocabulary and specific Education levels scores were not collected for the Younger Adults.

5.3.2 Materials

The experiment was programmed and run in PsychoPy (Peirce 2007). The same word stimuli were used as in Experiment 3, and these were organised and counterbalanced across conditions in the same way as in Experiment 3 (see Chapter 4). Milk chocolate buttons were provided as a consistent external reward to encourage task engagement and motivation, but reward was not manipulated as a factor.

As in the previous experiments and a set of 80 emotionally neutral images from the IAPS database (Lang et al., 2005) were used in a visual search filler task. The filler task was the same as described in Chapter 2 (2.3.3.2) with two repetitions of the task presented, once after the encoding and then again after the first recognition test.

5.3.3 Design and Procedure

The paradigm followed exactly the same protocol as in the pilot study. In summary, all participants first completed an encoding task, where they were instructed to intentionally learn words. Then, participants completed an old/new recognition task (Test 1) and finally a surprise

---

[6] The MoCA threshold of >=25 was used instead of the standard 26, in order to be less conservative with regards to inclusions. Since the behavioural results showed a lack of memory performance difference between the age groups, we did not want to restrict group differences by being overly conservative. For the record, all key conclusions from analyses were the same when a MoCA cut off of >=26 was used.

old/new recognition task for foils that had been shown in Test 1. Each surprise test old/new decision was directly followed by a source judgment for all ‘old’ responses, requiring participants to judge whether they remembered seeing the word in the study and test phases. All old/new recognition responses and source judgments were given with continuous confidence measures as used in previous studies. Like in previous experiments, in between each of the three phases, a visual search filler task was used to clear working memory. The experimental paradigm is summarised in Figure 5.1.

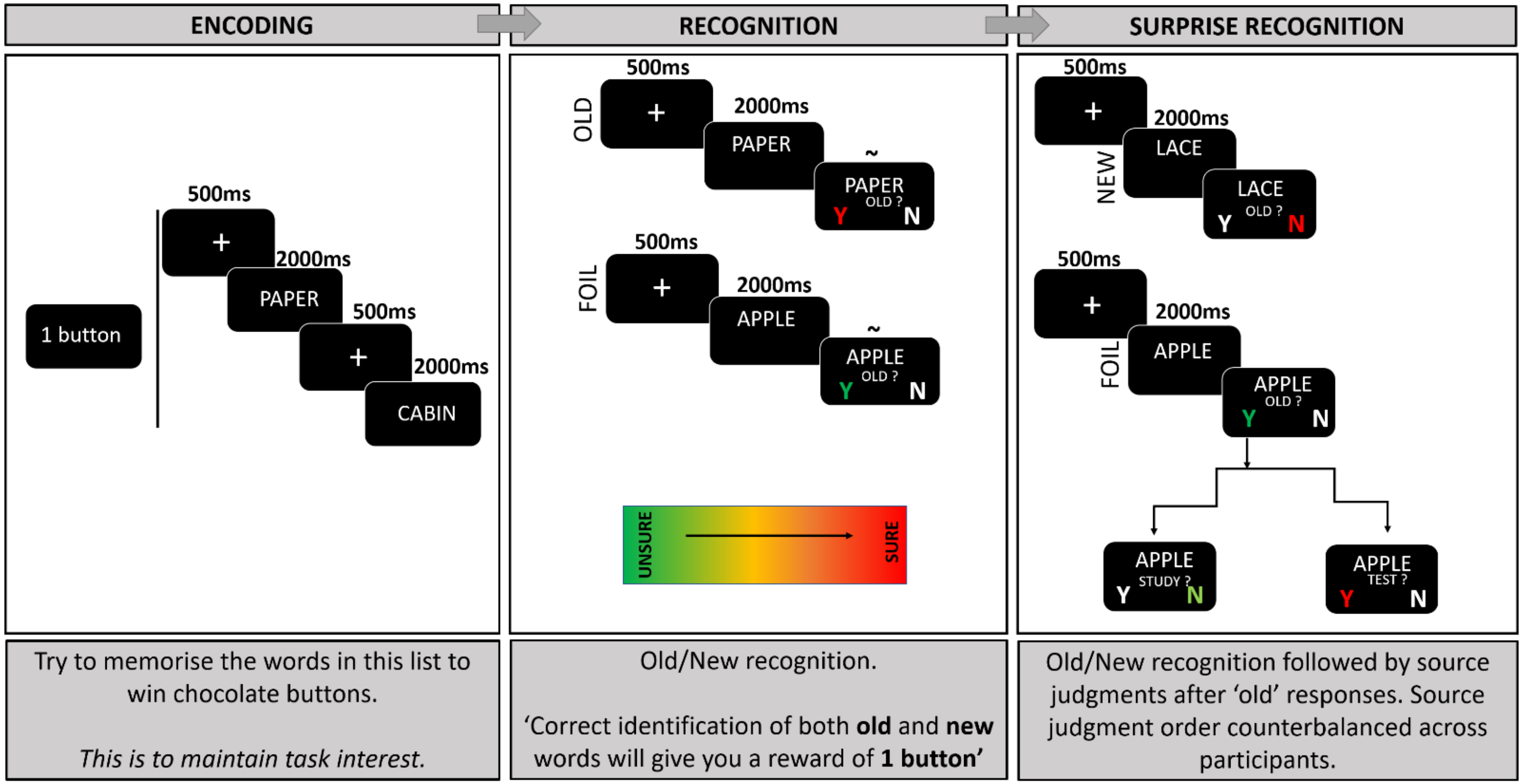


Figure 5.1. Schematic of Experiment 5 procedures. There were three phases, intentional encoding, recognition and a surprise recognition phase. Continuous confidence scores were made during the initial recognition tests and surprise tests. During the Surprise test, source judgments were made for words that were responded to as 'Old'.

During the encoding phase, participants were asked to intentionally try to learn a list of 128 words, which were shown for 2000ms after a 500ms fixation cross. The study phase was split into two halves, and words within each half were presented in a random order. There was a break screen mid-way after 64 encoding trials. Participants were told that they would win 1 chocolate button every time they correctly later identified a word in the upcoming recognition phase. This reward was included to try to keep the participants motivated and attentive throughout the lengthy study and test phases, but reward was not manipulated.

Next, participants proceeded to the old/new recognition test phase where they were asked if they recognised words from the previous study phase or not. Previously seen "old" words were intermixed with "new" foil words, and participants responded to whether each word was 'old' by pressing buttons on the keyboard. This test included 256 words in total, with 128 'old' words and 128 'new' items. Words that had been shown in the first half of the study task were tested first, and words that had been shown in the second half of the study phase were always tested after all the words from the first block, to ensure a similar study-test delay across items. A break was given after every 64 trials.

During each recognition test trial, a fixation cross was presented for 500ms before each word was shown in the middle of the screen for 2000ms. Participants were told to not respond until the response options were shown after 2000ms. The response options were presented with the prompt question 'Old?' and the two options 'Yes' and 'No'. Responses were made with the 'Z' and the 'M' key (button press measured RT). Allocation of these keys to response options was counterbalanced across participants. As in Experiment 3, participants were also asked to indicate their confidence in their recognition decision using the button press duration; the longer they held the key for the more confident they were in their recognition decision (button press duration: measured with key release time). Visual feedback was shown on the screen with a continuous colour change from vivid green (unsure) to vivid red (very sure) for their chosen option. Unlike in Experiment 3 where participants had a fixed time of 2700ms to make their decisions and indicate confidence, within this experiment participants had unlimited time to respond (as explained previously this change was implemented to remove a confound between response times and confidence). When they released the button, the next trial started with a fixation cross. There was an extensive practise session provided before the test to familiarise participants with the confidence scale and key responses.

Then participants were told about the surprise recognition test, where they were instructed that words would be shown on the screen and these could have appeared at any time during the experiment so far, or could be completely new. Participants were tasked to make a recognition judgment whether they had seen the word before at any point during the whole study. Overall, during the surprise recognition test participants saw 270 words in total. They were presented with all previously seen foil words from the recognition test 1 phase (128 foil words) as well as 10% of the studied items for the study phase (14 studied words, 7 from each half of the study phase) intermixed with completely new words (128). Participants were unaware of the proportion of words from these categories.

Words were shown for 2000ms, preceded by a fixation cross for 500ms. Participants used the 'Z' and 'M' keys to respond 'Yes' or 'No' to the question 'Is the word 'Old'?'. Key response mappings were counterbalanced across participants but kept consistent with previous tasks. Similarly to in Recognition test 1, participants also provided their confidence alongside their recognition responses by using their button press duration, indicated on the screen with a colour change (as described previously). Once again participants had unlimited time to respond during the surprise test. This judgment was previously not made in Experiment 3 but was added to provide a measure of item recognition confidence.

Another addition as compared to Experiment 3 was that if participants responded that the word was "old", two follow up source questions were asked; 'Did you see this word in the study phase?' and 'Did you see this word in the test phase?' (The order of these questions was counterbalanced across different participants, but kept constant within participants). Again, participants indicated 'Yes' or 'No' with the 'Z' and 'M' keys (with same response mapping as for the other judgements), and also indicated their confidence in the chosen response using the button press durations. Extensive instructions and practice were given to explain the source judgments and to ensure that participants understood the difference between each phase and how to respond to the source judgment questions. Participants were told that words could have been shown in the study or test phases or in both phases, and that therefore they should make each judgement independently (i.e. the options were not mutually exclusive).

In between the study, test 1 and surprise test experimental phases, a visual search task was used as a filler to clear working memory. The task was exactly the same as in prior studies and lasted for 1 minute in total (see Chapter 2 for detailed description).

Full counterbalancing of word assignment to conditions, response hand, and source question order was achieved across the first 24 participants in each group, and counterbalancing formats were assigned in a pseudo-random way to the remaining participants to balance all factors. This involved a systematically selecting a subset of each counterbalance rotation to be re-sampled ensuring as much as possible an equal rotation of stimuli, button press order and source question order across both Older and Younger participants.

After completing all computerised tasks, all participants were asked to complete a questionnaire about their motivation during the study and test phases, the effort they put into learning or recognising words, and if they expected the surprise test, as well as questions about any strategies or memory techniques used. For the older group only, additional questionnaires were collected; The Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005) and the Shipley Vocabulary Test (the Vocabulary subsection of the Shipley Living Scale; Carlozzi, 2011) were used to determine 'healthy' cognition and a vocabulary level, and an additional questionnaire collected demographic data such as education level, occupation, physical and social activity levels.

5.3.4 EEG recording and pre-processing

Continuous Electroencephalography (EEG) was recorded throughout all experimental tasks at 500 Hz with a 0.1-70Hz bandwidth, referenced against an average reference. Twenty-nine electrodes attached to an EasyCap (Brain Products Gmbh) recorded scalp EEG, with locations according to the extended 10-20 system. Impedances were kept below 10kΩ. The electrooculogram (EOG) was recorded vertically (VEOG) from the right eye and horizontally (HEOG) from the right and left eye.

The raw EEG data were pre-processed using the EEGLAB toolbox (UC San Diego, 2004) following a similar pipeline as described in Bergstrom, et al. (2016). The EEG data were re-referenced to the average of the mastoids and segmented into 526 epochs time-locked to the onset of the word in the two recognition tests (2000ms, including a 500ms pre-stimulus period, of which -200ms to 0ms was used for baseline correction). An initial digital filter with a high-pass of 0.1Hz was applied. Visual inspection of epochs was conducted to remove any extremely noisy trials, for example from coughing or large muscle movement. The remaining epochs were then concatenated and submitted to extended Infomax Independent Component Analysis (ICA) using 'Runica' from the EEGLAB toolbox, with default extended-mode training parameters (Delorme & Makeig, 2004). Independent components reflecting eye movements were identified by examining the output from both ICLabel (Delorme & Makeig, 2004) and ADJUST (Mognon, Jovicich, Bruzzone, & Buiatti, 2011), coupled with visual inspection of component scalp topographies, time courses, and activation spectra. Only components that were unambiguously identified as reflecting eye-movements were selected for removal to ensure equivalent and unbiased noise removal between age groups (ADJUST and ICLabel were used only as a guide and not as an automatic ICA removal process). These components were then discarded from the data by back-projecting all but these components to the data space. Corrected data was subsequently digitally lowpass filtered at 40 Hz. Epochs were then manually visually inspected again and any trials that still contained visible artefacts after filtering were removed, as were trials where RTs were over 10 seconds. This lax RT threshold was used to ensure there were no group differences in trial removal due to RTs whilst also removing any extreme RT outliers (in order to remove potential unusual brain-activity for example associated with attentional lapses). Overall, only a small percentage of trials were removed in each group (2.6% of Younger Adults' data and 3.3% of Older Adults' data). Only participants who had enough trials per critical condition (*N>14*) and who had sufficiently clean ERPs were included for further analysis.

5.3.5 Data Analysis Protocol

5.3.5.1 Behavioural Data analysis

For the Test 1 phase, the proportion of correct responses were calculated separately for Old words (i.e. hit rates) and New Foil words (i.e. correct rejection rates), and these were used to derive measures of discrimination (*Pr*) and response bias (*Br*), calculated as in prior experiments to enable comparison with prior findings (see Chapter 2, section 2.3.4 for details). As in Experiment 3, new foils were also separated into conditions based on participants' accurate (correct rejections, "CRs") or inaccurate (false alarms, "FAs") responses on a within-subjects basis. For all foils receiving CRs, the median RT for those trials within each participant was used to categorise trials into "Fast CRs" that were responded to faster than the median threshold, and “Slow CRs” than were responded to slower than the median threshold[7]. Mean RT and confidence were extracted to characterise the different conditions.

On the surprise foil recognition test, overall measures of discrimination (*Pr*) and response bias (*Br*) were calculated using all old words and all new words. Here, prior foils were sorted into conditions based on the previous responses that participants had given during Test 1, leading to three main conditions: “Prior FAs”, “Prior Fast CRs” and “Prior Slow CRs”. *Pr* was calculated for each of these foil conditions separately to assess how discrimination performance on the surprise test was modulated by prior responses on test 1, and whether that pattern differed across age groups. Because this *Pr* calculation used a constant FA rate (as there was only one surprise test and hence only one category of new items), any within-subject differences between conditions would be the same as if comparing raw hit rates, but importantly the *Pr* measure removes the influence of response bias on scores, which could otherwise potentially differ between age groups.

In addition to recognition accuracy, mean confidence in recognition judgements was also compared for the conditions described above, however RTs were not analysed on this test since participants were told to prioritise accurate rather than fast decision making. Furthermore, I also extracted and compared the proportion of recognised foils (i.e. foils that participants correctly recognised as previously seen) where participants thought they remembered seeing the word in the study phase and the test phase for the different foil conditions. Because none of the foils were actually shown during the study phase but all foils were shown during the test phase, these proportions represent errors to the “study” source question but correct responses to the “test” source question. The mean confidence for source judgements was also compared across foil conditions.

---

[7] Note that a similar analysis was conducted for confidence, but since confidence and RTs were strongly correlated this analysis was redundant and is therefore not presented.

The above measures were compared across conditions using a combination of frequentist inferential statistical tests coupled with Bayes Factors, as in previous experiments (for detailed description see Chapter 2, section 2.3.4).

5.3.5.2 EEG Data Analysis Protocol

ERPs were formed for key conditions of interest separately for each memory test (Recognition Test 1 and Surprise Recognition Test) and age group. The mean trial numbers per condition can be seen in Table 5.1. Trials could not be split by source accuracy due to low trial numbers for False Alarm responses on the Surprise Test. Including only participants who had sufficient trials for this particular analysis would have potentially introduced a confound into the type of False Alarms being investigated. Those with more False Alarms are more likely to rely on guesses and gist judgments rather than true false memories for items. Therefore, I decided to only use source judgment responses in an informative way, and not in the following EEG analysis. Similarly, trial numbers mean that any analysis into DM effects of subsequent memory could not be conducted either. I discuss in the General Discussion suggestions to increase trail numbers.

Table 5.1. *Trial numbers included in the EEG analysis split by Age Group and Condition.*

Test Phase	Word Type	Condition	Younger Adults	Older Adults
Test 1	Old	Hit	95 (74-118)	91 (50-125)
	New Foil	CR	84 (59-113)	91 (54-111)
		FA	41 (15-68)	33 (14-66)
		Fast CR	42 (29-56)	46 (27-55)
		Slow CR	42 (30-57)	46 (27-57)
Surprise Test	Foil	Prior FA: Hit	38 (14-56)	30 (14-55)
		Prior Fast CR: Hit	38 (21-55)	35 (19-52)
		Prior Slow CR: Hit	38 (22-55)	37 (21-54)
		All Foils: Hit	82 (80-127)	79 (73-124)
	New	CR	82 (48-114)	79 (48-104)

Note: Values are mean trial numbers and Min-Max in brackets.

First the pre-registered predictions were tested by statistically analysing ERP mean amplitudes from two a-priori selected time-windows using a grid of four electrode sites that spanned the predicted scalp locations of effects (left and right frontal sites; F3 and F4, and the left and right parietal sites; P3 and P4). Mean ERP amplitudes were extracted between 300-500ms post-stimulus to test for FN400 old/new effects and between 500-800ms post-stimulus to test for left parietal old/new effects. These pre-registered predictions were based on the typical scalp distributions and timing of recognition-related ERP effects in young adults (Rugg & Curran, 2007). I used omnibus ANOVAs within each relevant time window to examine pairwise ERP differences between experimental conditions, whether those differences interacted with age groups, and whether they had different scalp topography across age groups. These omnibus ANOVAs therefore included factors for Condition (*'C'*: two levels that differed depending on which pairwise conditions were compared), Anterior-Posterior location (*'AP'*: Frontal vs. Parietal), Hemisphere (*'H'*: Left vs. Right) and the between groups factor; Age Group (*'Age'*: Younger Adults vs. Older Adults). Since only main effects of Condition or interactions between the other factors with Condition (C) are functionally meaningful, only those effects were followed up on and reported. If interactions between Age group or location factors with Condition were significant, the highest level of significant interaction was always followed up on. Follow-up tests were done for the highest level of interaction, for example into here was a 4-way interaction (Age x C x AP x H), this was followed up with two 3-way analyses separately for each age group (C x AP x H). If a 3-way interaction was subsequently seen, this was then directly tested with pairwise t-tests at each electrode site to investigate the location of the effect as this was the most parsimonious way to investigate the 3-way interaction with Condition, Anterior-Posterior location and Hemisphere.

Selecting only specific targeted time-windows and electrode sites for analysis may overlook effects at other sites or outside of these time-windows. This is problematic when comparing older and younger groups who may have different scalp distribution or timing of effects, and when recording ERPs in novel paradigms (such as in the current study) where the scalp distribution and timing of effects are not known a priori. Therefore, I also performed whole-scalp analysis to supplement the targeted analysis. Nonparametric cluster-based permutation tests were used as a data-driven, whole scalp method (FieldTrip toolbox; Oostenveld, Fries, Maris, & Schoffelen, 2011) to explore within and between group effects with minimal assumptions about time-windows and distributions of potential effects, whilst controlling for multiple comparisons (which was not the case for the targeted ERP analysis).

In the first step, t-tests were performed at every ERP data point to test for significant differences between conditions and groups (with uncorrected α=.05). Cluster groupings were next formed where t-values that were significant across adjacent electrodes (minimum of 2 electrodes) were summed to create a cluster-level *t*-value. This observed cluster-level sum t-value was then compared against a null distribution created by permutation resampling to control for inflated false positive rates from the vast number of initial t-tests. The null distribution was created by randomly permuting the conditions across participants (5000 times) and then testing for significant clusters using the same steps as described above, thereby creating a null distribution of summed cluster t-values. The significance (*p*-value) of each observed cluster was calculated as the proportion on the null distribution that was larger than the observed cluster-level test statistic (i.e. the Monte-Carlo *p*-value). This enabled me to identify which observed clusters were significant against a corrected α=.05 (two-tailed) and characterise how they extended over time and space (see Maris & Oostenveld, 2007). Using this complementary analysis, I tested if there were ERP differences between the same pairwise conditions as in the targeted analysis, both within each age group and I also tested if the ERP condition differences interacted with age group.

5.4 Experiment 5 Results

Additional Bayesian outputs can be found on the Open Science Framework website along with additional Bayesian outputs with supplementary materials being found in Appendix 4.1.2 and 4.2.2 (https://osf.io/w4x6h/?view_only=4e0b83becdbb4e919352b7a55d8a06ad).

5.4.1 Behavioural Analysis of all trials

Because the EEG analysis required removal participants who had too few trials per condition for ERPs and also delete noisy trials or trials with extreme RTs, this could have produced different behavioural results in the final sample to those found in my prior experiments. Therefore, I first conducted a behavioural analysis including a larger sample (*N*=79) and using trial inclusion criteria consistent with the prior behavioural studies (a laxer criterion on required trial numbers for inclusion in analyses; *N*=10 instead of *N*=14 which was used for EEG inclusion). This analysis was conducted to test if participant's responses to foils on the first test influenced subsequent Surprise Test foil recognition in the same way as in my prior experiments, and whether this influence differed across age groups.

The full results of this analysis are described in Appendix 4.2.2.1, but in summary, the main finding was that foils that received FAs on Test 1 were recognised more often as "old" on the surprise test compared to foils that received CRs on Test 1, with prior Slow CRs being recognised least often of all, thus replicating the pattern in Experiment 3 and pilot experiment. There was a larger effect in the Older Adults, who showed differences between all foil conditions on the surprise test (prior FAs> prior Fast CRs> prior Slow CRs) whereas Younger Adults recognised both prior FAs and prior Fast CRs better than prior Slow CRs (prior FAs &

prior Fast CRs > prior Slow CRs). Source responses on the surprise test also replicated findings from the pilot study, with participants more likely to show incorrect source memory for having seen foils in the study phase if those foils had received FAs on the first test, with Younger Adults making more such source errors than Older Adults.

5.4.2 Behavioural Analysis of EEG trials

Next, I conducted a behavioural analysis using a stricter trial inclusion criterion (N>14, as required for EEG) and only including the exact same trials as used in the ERP analysis (thus after EEG pre-processing trial rejection, see Methods section). Therefore, the behavioural data here is directly linked to the neural data analysed and is presented in order to facilitate interpretation of ERP effects, although it should be noted that selective exclusion of participants with particular performance patterns is likely to have biased the behavioural results (see discussion for further explanation on this point).

5.4.2.1 Recognition Test 1

Discrimination (*Pr*) and Response Bias (*Br*) were calculated to provide a comparable measure of recognition performance as used in previous Experiments. *Pr* and *Br* scores can be seen in Table 5.2 along with raw recognition accuracy scores. There was no difference between Younger and Older Adults in their ability to discriminate between old and new items on Test 1 (*Pr*; $t(60)= 1.06$, $p=.291$, $d=.27$; $BF_{10}=1/2.40$), and no Age Group difference in Response Bias (*Br*) on Test 1 ($t(60)= 1.69$, $p=.097$, $d=.42$; $BF_{10}=1/1.18$), and the Bayes Factors were anecdotally more consistent with no difference than a difference between groups for both measures.

Table 5.2. *Performance on the first recognition test. The table shows proportion accurate responses, Discrimination (Pr)and Response bias (Br) split by age group.*

Age Group	Recognition Accuracy				Discrimination (*Pr*)		Response Bias (*Br*)	
	Old Words		New Words					
Younger Adults	0.76	(0.09)	0.67	(0.11)	0.42	(0.16)	0.58	(0.13)
Older Adults	0.73	(0.13)	0.73	(0.11)	0.47	(0.14)	0.51	(0.17)

Note: Values shown are Means (SD).

Correct rejection responses to foils on Test 1 were split on a within-subjects basis into Fast vs. Slow CRs (RTs below or above the median CR RT for each person, as in Experiment 3), and these conditions were compared to foils that received FAs on Test 1 in order to investigate potential novelty detection vs. familiarity effects on ERPs and memory performance[8]. Mean RT and confidence for these foil conditions on Test 1 are summarised in Table 5.3, where it can be seen that for both Younger and Older Age Groups, RTs and Confidence for the FA condition fell in between the Fast and Slow CRs conditions, similarly to the pattern seen in Young Adults in Experiment 3 (see Appendix 4.1.2 for pilot study results). That is, for both Age Groups, FA responses to foils on Test 1 were associated with faster RTs than the Slow CR condition but slower RTs than the Fast CR condition, and were associated with higher confidence than the Slow CR condition but lower confidence than the Fast CR condition.

---

[8] Additional median split analyses were conducted based on Test 1 Confidence (high confidence CRs vs. low confidence CRs), but because confidence and RTs were highly correlated these analyses were redundant, so the confidence split analysis is presented in Appendix 4.2.2.2.

Table 5.3. *Response times and Confidence for foil conditions on Test 1, separately for Younger and Older groups.*

Measure	Age Group	FA		Fast CR		Slow CR	
Response Time	Younger Adults	591	(329)	333	(142)	966	(564)
	Older Adults	702	(639)	275	(259)	917	(640)
Confidence	Younger Adults	742	(392)	1214	(431)	737	(440)
	Older Adults	602	(497)	1016	(743)	566	(450)

Note: Values are means (*SD*) of Response Times and Confidence in *ms*.

5.4.2.2 Surprise Foil Test Phase

Old/New recognition judgements

Overall Discrimination (*Pr*) and Response Bias (*Br*) on the Surprise Test was first used to test for differences in general surprise test recognition performance between groups. Overall *Pr* and *Br* were calculated by collapsing across all prior Foil Conditions and all New Words, using the same formula as in previous chapters. As seen in Table 5.4, in contrast to Test 1 there was a significant difference in Discrimination ability (*Pr*) between Age Groups on the surprise test, with Younger Adults showing better discrimination than Older Adults ($t(54.45)= 3.32, p=.002, d= 0.85; BF_{10} = 21.43$), which was supported by strong evidence from the Bayesian analysis for the alternative hypotehsis. Additionally, Younger Adults also had a more liberal Response Bias (*Br*) than Older Adults ($t(60)= 2.32, p=.024, d= 0.59; BF_{10} =2.36$) with Bayesian analysis showing moderate support, and it should be noted that both groups were more liberal than conservative in their response bias, tending to respond ‘Old’ when uncertain.

Table 5.4. *Performance on the Surprise recognition test. The table shows proportion accurate responses, Discrimination (Pr)and Response bias (Br) split by age group.*

Age Group	Recognition Accuracy		Discrimination (*Pr*)	Response Bias (*Br*)
	Old Words (Prior Foils)	New Words		
Younger Adults	0.91 (0.07)	0.66 (0.14)	0.57 (0.15)	0.77 (0.15)
Older Adults	0.82 (0.11)	0.64 (0.13)	0.46 (0.10)	0.67 (0.20)

Note: Values shown are Means (*SD*) for all types of prior foils combined and completely new items.

The next analysis was conducted for the separate foil conditions to test the prediction that there would be a higher level of subsequent recognition for foils that were previously incorrectly identified as old on Test 1 (prior FAs) compared to foils that were previously correctly identified as new on Test 1 (prior Fast CRs and Slow CRs), and that prior Fast CRs would be more often recognised as old than prior Slow CRs. Discrimination scores (*Pr*) for these conditions are shown in Table 5.5 along with associated Confidence in decisions.

Table 5.5. *Discrimination ability and confidence on the Surprise Test split by prior Foil Condition and Age group.*

Measure	Age Group	Prior Foil Condition		
		Prior FA	Prior Fast CR	Prior Slow CR
Surprise Test Discrimination (*Pr*)	Younger Adults	0.60 (0.14)	0.61 (0.15)	0.59 (0.14)
	Older Adults	0.58 (0.12)	0.50 (0.10)	0.52 (0.09)
Confidence (*ms*)	Younger Adults	1440 (317)	1481 (324)	1455 (326)
	Older Adults	1220 (833)	1240 (897)	1168 (920)

Note: Values shown are Means (SD). Confidence shown in *ms.*

Separate mixed ANOVAs tested for the predicted pairwise differences in surprise test discrimination (*Pr*) between the Prior Foil Conditions, including Age Group (Younger vs. Older) as a factor to test whether foil differences interacted with age (Table 5.6).

Table 5.6. *The results of Mixed 2x2ANOVAs, assessing pairwise effects of Prior Foil Condition x Age Group (Younger/ Older Adults) on Discrimination ability (Pr) on the Surprise Test.*

Prior Foil Conditions	Effect	*F*	*p*	*np2*
FA x Fast CR	Word Type	**27.67**	**<.001**	**.32**
	Age	**4.97**	**.03**	**.08**
	Word Type *Age	**27.67**	**<.001**	**.38**
FA x Slow CR	Word Type	**22.16**	**<.001**	**.27**
	Age	2.48	.121	.04
	Word Type *Age	**11.34**	**.001**	**.16**
Fast CR x Slow CR	Word Type	1.17	.285	.02
	Age	**9.25**	**.003**	**.13**
	Word Type *Age	**13.67**	**<.001**	**.19**

Note. Significant values below .05 are shown in bold. *N*= 62

These mixed ANOVAs showed that surprise test *Pr* was higher for prior FAs than for prior Fast CRs as predicted, but this effect also interacted with Age Group. Separate analysis within each Age Group revealed that the Older Adults were better able to discriminate prior FAs than prior Fast CRs from new items (*t*((29)=6.75, *p*<.001, *d*= 0.81; BF_{10} = 75179.23) supported by extreme Bayesian evidence for the alternative hypothesis, but there was no such difference in discrimination for Younger Adults (*t*(31)= 0.65, *p*=.523, *d*= 0.04; BF_{10}=1/4.37) supported by moderate Bayesian evidence for the null hypothesis here. Surprise test *Pr* was also higher for prior FAs than for prior Slow CRs as predicted, but this effect also interacted with Age Group. This interaction was similarly driven by higher *Pr* for prior FAs than prior Slow CRs within the Older Adults only (*t*((29)=5.46, *p*<.001, *d*=0.57; BF_{10}=2921.01) and no such effect in the Younger Adults *t*(31)= 0.99, *p*=.330, *d*=0.07; BF_{10}=1/3.38). When comparing prior Fast CRs to prior Slow CRs, there was no main effect of prior foil type on surprise test *Pr,* however there was a significant interaction between foil type and Age Group. When examining this interaction within each Age Group, there were opposing directional effects. Older Adults were significantly better at recognising prior Slow CRs than prior Fast CRs (*t*((29)=3.10, *p*=.004, *d*=0.30; BF_{10}=9.34). Whereas within the Younger Adults, there was a trend towards an opposing effect, with prior Fast CRs being recognised slightly more than prior Slow CRs (*t*(31)= 2.03, *p*=.051, *d*= 0.11; BF_{10}=1.14), in line with the findings in Experiment 3.

Similar mixed 2x2 ANOVAs (Table 5.7) tested for pairwise differences in recognition judgement confidence between prior foil types and whether those differences interacted with Age Group (Younger vs. Older). However, there were no significant effects of prior foil condition or age on recognition confidence in the surprise test.

Table 5.7. *The results of mixed 2x2ANOVAs, assessing pairwise effects of Prior Foil Condition x Age Group (Younger/ Older Adults) on Recognition Confidence on the Surprise Test.*

Prior Foil Conditions	Effect	*F*	*p*	*np2*
FA x Fast CR	Word Type	1.47	.231	.02
	Age	2.03	.160	<.01
	Word Type *Age	0.16	.687	<.01
FA x Slow CR	Word Type	0.38	.542	<.01
	Age	2.42	.125	.04
	Word Type *Age	1.21	.276	.02
Fast CR x Slow CR	Word Type	3.84	.055	.06
	Age	2.43	.124	.04
	Word Type *Age	0.87	.355	.01

Note. Significant values below .05 are shown in bold. *N*= 62

5.4.2.3 Source Judgments

To recap, source judgments were made in response to two questions; “Was the word shown in the ‘Study’ phase?” and “Was the word shown in the ‘Test’ phase?”. Proportions ‘Yes’ responses for each question (Table 5.8) were compared across prior foil conditions and age groups (note that this measure therefore indicates inaccurate source memory for the study phase, but accurate source memory the Test 1 phase).

Table 5.8. *Proportion of 'Yes' responses to the two Source Questions split by prior Foil type and Age Group.*

Source Question	Age Group	Prior Foil Conditions		
		FA	Fast CRs	Slow CRs
Remember from Study?	Younger Adults	0.60 (0.19)	0.29 (0.27)	0.34 (0.24)
	Older Adults	0.45 (0.21)	0.24 (0.23)	0.24 (0.24)
Remember from Test?	Younger Adults	0.83 (0.19)	0.86 (0.23)	0.83 (0.22)
	Older Adults	0.80 (0.14)	0.83 (0.19)	0.85 (0.19)

Note: Values shown are Means (SD).

Mixed 2x2 ANOVAs were again used to test for pairwise differences in source memory between prior Foil Conditions as well as testing for differences or interactions with Age Group (Table 5.9). Differences in source memory between conditions and Age Groups were seen only for the incorrect responses to the 'Study?' source question, with no significant differences in any comparisons for correct responses to the 'Test?' source question. Participants were more likely to report false source memory for having seen foil words during the Study phase for prior FAs than both prior Fast CR and for prior Slow CRs. Neither of these differences interacted significantly with Age Group, but there was a main effect of Age Group for the latter comparison, with Younger Adults reporting more incorrect source memory for the study phase than Older Adults overall. There were no differences in incorrect source memory for the study phase when comparing prior Fast CRs and prior Slow CRs.

Table 5.9. *The results of Mixed 2x2ANOVAs assessing the effects of prior foil type and age group on source memory for having seen foils in the study and test phases.*

Measure	Prior Foil Conditions	Effect	*F*	*p*	*np2*
Responding 'Yes' to the Source Question: "*Study?*"	FA x Fast CR	Word Type	**112.72**	**<.001**	**.65**
		Age	3.78	.057	.06
		Word Type *Age	3.78	.058	.06
	FA x Slow CR	Word Type	**127.24**	**<.001**	**.68**
		Age	**5.51**	**.022**	**.08**
		Word Type *Age	1.56	.217	.03
	Fast CR x Slow CR	Word Type	3.48	.067	.06
		Age	1.57	.217	.03
		Word Type *Age	2.23	.141	.04
Responding 'Yes' to the Source Question: "*Test?*"	FA x Fast CR	Word Type	1.00	.320	.02
		Age	0.10	.754	<.01
		Word Type *Age	0.17	.685	<.01
	FA x Slow CR	Word Type	0.51	.478	.01
		Age	0.02	.899	<.01
		Word Type *Age	0.97	.329	.02
	Fast CR x Slow CR	Word Type	0.59	.444	.01
		Age	**6.61**	**.013**	**.10**
		Word Type *Age	**6.60**	**.013**	**.10**

Note. Significant values below .05 are shown in bold. *N*=62

Confidence in Study Source Responses

Confidence ratings were provided for each Source Question separately, and these were compared across foil type conditions and age groups (Table 5.10).

Table 5.10. *Mean confidence in source memory judgements regarding study and test phases for the different foil types and Age Groups.*

Source Question	Age Group	Prior Foil Conditions		
		FA	Fast CRs	Slow CRs
Source Question: Study	Younger Adults	979 (382)	1136 (424)	1081 (382)
	Older Adults	673 (552)	721 (639)	612 (575)
Source Question: Test	Younger Adults	1353 (382)	1498 (328)	1413 (335)
	Older Adults	938 (817)	936 (787)	1035 (838)

Note: Values are the Mean (*SD*) Confidence (in *ms*) of subject-level median scores for each condition.

The mixed ANOVAs (Table 5.11) that compared source memory for the study phase between pairwise Foil Conditions and Age Groups showed that Younger adults were overall more confident in their Study phase source judgements than older adults, across all three pairwise ANOVAs. Prior Fast CRs were associated with more confident source judgements about the study phase than Prior FAs and Prior Slow CRs, and neither of these foil condition differences interacted with Age Group. For the ANOVA comparing Prior FAs and Prior Slow CRs, there was a significant interaction between this Foil Condition difference and Age Group however. Follow up tests showed that Younger Adults were significantly more confident in the Study source responses for prior Slow CRs than for prior FAs ($t(31)=2.24$, $p=.032$, $d=0.27$; $BF_{10}=1.65$), whereas Older Adults did not differ in their confidence between these foil types ($t(29)=0.94$, $p=.356$, $d=0.11$; $BF_{10}=1/3.44$) with moderate Bayesian evidence for the null hypothesis.

Table 5.11. *The results of Mixed 2x2ANOVAs assessing effects of foil type condition and age group on confidence in source judgements about the study and test phases.*

Measure	Prior Foil Conditions	Effect	*F*	*p*	*np2*
Confidence in Source Question: *Study?*	FA x Fast CR	Word Type	**4.70**	**.034**	**.07**
		Age	**9.09**	**.004**	**.13**
		Word Type *Age	1.29	.260	.02
	FA x Slow CR	Word Type	0.29	.591	.01
		Age	**11.33**	**.001**	**.16**
		Word Type *Age	**4.34**	**.042**	**.08**
	Fast CR x Slow CR	Word Type	**7.82**	**.007**	**.12**
		Age	**12.11**	**.001**	**.17**
		Word Type *Age	0.86	.358	.01
Confidence in Source Question: *Test?*	FA x Fast CR	Word Type	**7.64**	**.008**	**.11**
		Age	**10.10**	**.002**	**.14**
		Word Type *Age	**8.08**	**.006**	**.12**
	FA x Slow CR	Word Type	**5.59**	**.021**	**.09**
		Age	**6.39**	**.014**	**.10**
		Word Type *Age	0.31	.578	.01
	Fast CR x Slow CR	Word Type	0.45	.833	<.01
		Age	**9.53**	**.003**	**.14**
		Word Type *Age	**7.80**	**.007**	**.12**

Note. Significant p-values below .05 are shown in bold. *N*=62

The mixed ANOVAs that compared source memory for the test phase between pairwise prior Foil Conditions and Age Groups (also Table 5.11), showed that Younger adults were also overall more confident in their Test phase Source Judgements than Older Adults, across all three pairwise ANOVAs. Participants were less confident about their Test phase source memory judgements for prior FAs than prior Slow CRs, and this difference did not interact with Age Group. Participants were also less confident overall for prior FAs than prior Fast CRs but this difference did interact with age group. The interaction was caused by Younger Adults reporting significant lower confidence in their Test phase source judgements for prior FAs than prior Fast CRs ($t(31)=3.66, p<.001, d= 0.41; BF_{10}=34.53$) supported by very strong Bayesian evidence for the alternative hypothesis, whereas the Older Adults reported highly similar levels of confidence in Source Judgements for both types of foils ($t(29)=0.06, p=.953, d<.001; BF_{10}$

=1/5.14) supported by supported by moderate Bayesian evidence for the null hypothesis. When comparing the prior Fast CRs and prior Slow CRs Foil Conditions, there was again a highly similar interaction between Foil Type and Age Group. When the difference between foils was examined separately for each Age Group, Younger Adults reported higher confidence for prior Fast CRs than prior Slow CRs ($t(31)=2.79$, $p=.009$, $d=0.26$; $BF_{10}=4.83$), whereas Older Adults again reported similar levels of confidence in source judgements for both types of foils ($t(29)=1.66$, $p=.105$, $d=0.12$; $BF_{10}=1.53$), with the younger adults having moderate Bayesian evidence for the alternative hypothesis and the older adults only having anecdotal evidence.

5.4.2.4 Behavioural Results Summary

In summary, the behavioural results from Test 1 showed very similar recognition performance across age groups after an intentional encoding task. Age group differences in recognition were seen however when memory for incidentally encoded foils was tested with a surprise recognition test. In this test, younger adults showed overall better discrimination ability compared to older adults, and even though both groups showed a response bias toward guessing that items were ‘Old’ when uncertain, the younger adults were the most liberal. When memory-for-foils on the surprise test was compared as a function of prior recognition accuracy and RTs of recognition judgements on the first test, the main findings were that younger adults did not show any significant differences between prior FAs, prior Fast CRs and prior Slow CRs in foil recognition, but they were more likely to incorrectly judge that prior FAs had been shown in the study phase than either prior Fast or Slow CRs. The younger adults also showed strong differences in judgement confidence consistent with source confusion for prior FAs. The older adults showed increased subsequent recognition of prior FAs compared to both prior Fast CRs and prior Slow CRs. Surprisingly, they also showed an increase in subsequent recognition for prior Slow compared to prior Fast CRs, thus contrary to the result in Experiment 3 and the trend in the young group in the current study, who showed better recognition memory for prior Fast CRs than prior Slow CRs when examining the full behavioural sample. Older adults made less source errors to the 'Study' question compared to Younger Adults, but similarly were more likely to falsely report remembering prior FAs from the study phase compared to the other foil conditions. These results therefore largely replicate my prior findings in Experiment 3, and suggest that false recognition of foils on test 1 "carries over" to the surprise test in terms of false source memory for having seen those foils in the study phase.

5.4.3 ERP Results

For each pairwise condition comparison, I first present the targeted analysis that was conducted to test pre-registered hypotheses (https://osf.io/s2drf) for young adults. For all these analyses, I added Age Group as a between-subjects factor to test if the predicted ERP effects were modified by ageing. Each targeted analysis is complemented with a global, whole scalp analysis testing for pairwise condition differences without focusing on specific time-windows and electrode sites, and using cluster-based permutation tests to correct for false positives.

5.4.3.1 Recognition Test 1 ERP results

Test 1 Correct Recognition ERP effects (Hit v CR)

The first analysis compared ERPs for Old items that were correctly recognised (Hits) versus New items that were correctly identified as new (All CRs, collapsed across fast and slow responses), to test for typical familiarity and recollection-related ERP effects and investigate how these might differ between Age Groups. Grand-average ERPs from the four electrode sites for each condition are shown in Figure 5.2.

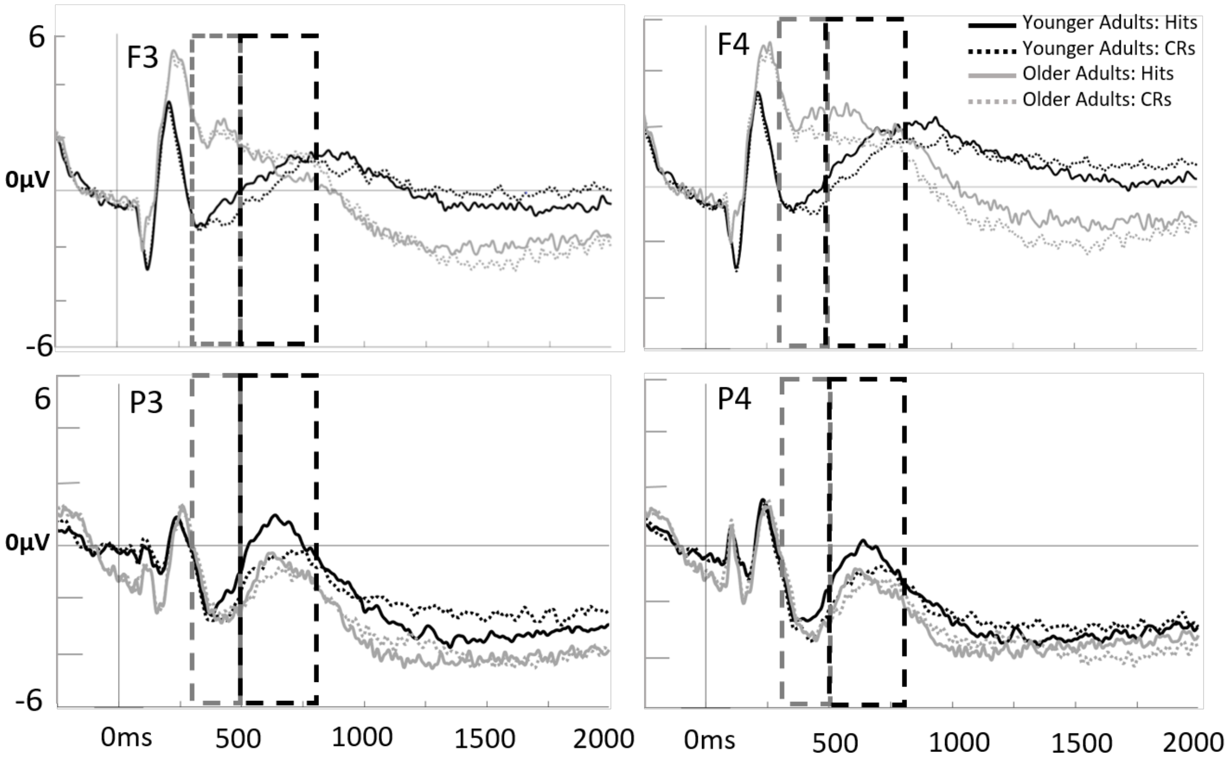


Figure 5.2. Grand-average ERPs for Hits and CRs in Test 1 for Younger and Older Adults at left and right frontal (F3, F4) and left and right parietal (P3, P4) electrode sites. Boxes highlight the early 300-500ms and the late 500-800ms targeted time windows.

Mixed omnibus ANOVAs (factors: Condition (Hits/CRs), Anterior/Posterior (Frontal/Parietal sites), Hemisphere (Left/Right sites), and Age Group (Older/Younger)) were conducted on mean amplitudes in early (300-500ms) and late (500-800ms) time windows to test for FN400 (associated with familiarity) and left parietal (associated with recollection) old/new effects respectively. These showed a marginally significant 4-way interaction in the early time window and a highly significant 4-way interaction in the late time-window (Table 5.12).

Table 5.12. *Results from a mixed ANOVA for the omnibus test comparing Age effects on ERPs associated with Hits vs. CRs during Test 1 in the two a-priori targeted time windows.*

Omnibus	Time Window					
	300-500ms			500-800ms		
	F	*p*	n_p^2	*F*	*p*	n_p^2
C	**4.98**	**0.029**	**0.08**	**15.04**	**<.001**	**0.20**
C x AP	1.88	0.176	0.03	1.70	0.198	0.03
C x H	2.61	0.112	0.04	**6.51**	**0.013**	**0.10**
C x AP x H	0.57	0.455	0.01	**41.05**	**<.001**	**0.41**
Age x C	3.51	0.066	0.06	**5.26**	**0.025**	**0.08**
Age x C x AP	**4.46**	**0.039**	**0.07**	0.23	0.636	<.001
Age x C x H	**4.78**	**0.033**	**0.07**	**10.45**	**0.002**	**0.15**
Age x C x AP x H	**3.79**	**0.056**	**0.06**	**10.43**	**0.002**	**0.15**

Note: C= Condition; AP =Anterior-Posterior; H = Hemisphere; Age = Age Group. Significant results are in bold. Only effects that involve the condition factor are reported. N=62

For completeness, I conducted follow up 3-way ANOVAs separately within each Age Group for both time windows, which revealed different patterns of effects in the two Age Groups (Table 5.13). In the early time-window, Hits were associated with more positive ERPs than CRs for Younger Adults, but this effect did not differ depending on hemisphere or anteriority of the electrode locations. In the later time-window for Younger Adults there was a 3-way interaction, whereas, within the Older Adults, there was a significant 3-way interaction between condition and anterior-posterior and hemisphere factors in both time windows. These significant interactions in the Older group and the late significant interaction in the Younger group were followed up by examining amplitude differences between Hits and CRs for each electrode site and time window separately (Table 5.14).

Table 5.13. *Results from the within group ANOVAs, testing ERP differences between Hits and CRs in the two selected time windows within Younger and Older groups separately.*

Condition Comparison	Age Group	Effect	Time Window					
			300-500ms			500-800ms		
			F	*p*	n_p^2	*F*	*p*	n_p^2
Hit v CR	Younger Adults	C	**6.81**	**0.014**	**0.18**	**17.47**	**<.001**	**0.36**
		C x AP	0.21	0.648	0.01	1.20	0.282	0.04
		C x H	0.13	0.717	<.01	0.35	0.559	0.01
		C x AP x H	0.58	0.454	0.02	**5.93**	**0.021**	**0.16**
	Older Adults	C	0.09	0.768	<.01	1.40	0.246	0.05
		C x AP	**9.20**	**0.005**	**0.24**	0.54	0.468	0.02
		C x H	**9.83**	**0.004**	**0.25**	**12.13**	**0.002**	**0.29**
		C x AP x H	**5.03**	**0.033**	**0.15**	**39.86**	**<.001**	**0.58**

Note: C= Condition; AP =Anterior-Posterior; H = Hemisphere; Age = Age Group. Significant results are in bold. Only effects that involve the condition factor are reported. N=62

In the early time window, Older Adults showed only a left frontal positivity for Hits compared to CRs. This finding thus suggested a more focal left frontal FN400 effect in Older Adults compared to the broadly distributed FN400 effect seen in Younger Adults. In the later time window, Older Adults had significant ERP differences between Hits and CRs at both frontal sites, but with reversed polarity. Whereas there was an increased positivity for Hits compared to CRs at the frontal right electrode (F4), CRs were more positive than Hits at the left frontal electrode (F3). Thus, Older Adults showed a frontal, right lateralised old>new effect in the 500-800ms time window when recollection-related ERP effects are typically found. The Younger Adults however showed a highly significant widespread old>new effect across all four electrode sites, but the effect was strongest at the left parietal site as typical for young adults in prior literature.

Table 5.14. *Results from follow up t-tests comparing Hits and. CRs at each electrode site for both the early and late time-window (500-800ms) within the Younger and Older groups.*

Condition Comparison	Electrode	Time Window							
		300-500ms				500-800ms			
		t	*p*	*d*	BF_{10}	*t*	*p*	*d*	BF_{10}
T1: Hit v CR; within Younger Adults	F3	-	-	-	-	**2.45**	**0.020**	**0.20**	**2.44**
	F4	-	-	-	-	**3.12**	**0.004**	**0.24**	**9.89**
	P3	-	-	-	-	**4.02**	**<.001**	**0.44**	**83.39**
	P4	-	-	-	-	**3.04**	**0.005**	**0.37**	**649.85**
T1: Hit v CR; within Older Adults	F3	0.52	0.608	0.02	1/4.54	**-2.62**	**0.014**	**-0.17**	**3.43**
	F4	**2.18**	**0.038**	**0.13**	**1.50**	**3.30**	**0.003**	**0.23**	**14.43**
	P3	-1.69	0.102	-0.09	1/1.46	1.62	0.115	0.12	1/1.60
	P4	-0.87	0.393	-0.04	1/5.95	0.86	0.397	0.06	1/3.66

Note: Significant results are in bold. Younger Adults N=32, Older Adults N=30

Consistent with the targeted analysis, the cluster-based permutation analysis (Fig. 5.3) indicated highly significant ERP differences between Hits and CRs within the Younger Adults. An initial cluster showed that ERPs for Hits were significantly more positive than CRs (p<.001) early in the epoch between approximately 410-730ms maximal across central and parietal electrodes, which was followed by a later left posterior negative cluster as Hits were less positive than CRs (p=.022) approximately between 1270-1440ms. Within the Older Adults, a similar early positivity was observed for Hits compared to CRs (p=.021), but this effect had a right lateralised topography between approximately 530-650ms. There was no significant late cluster in the Older adults, but ERP amplitudes were in the opposite direction in the late epoch in Older Adults compared to the Younger group, with CRs being less positive than Hits.

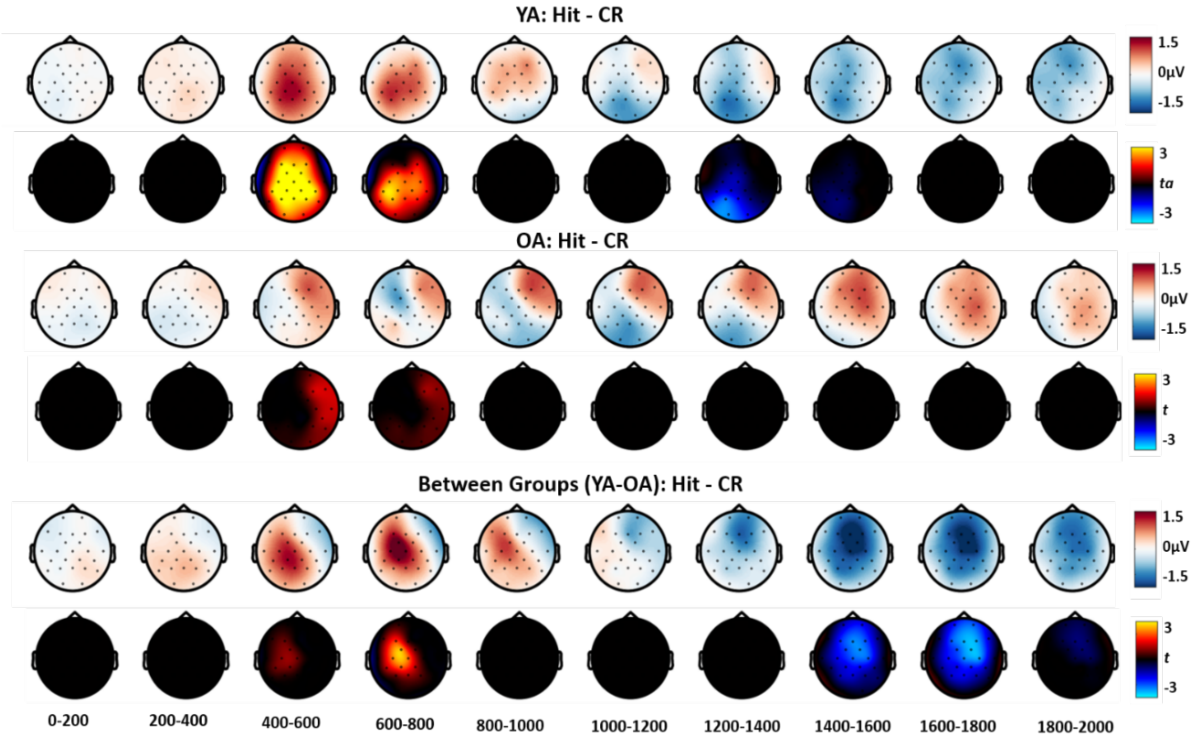


Figure 5.3. Whole-scalp analysis of Test 1 Hit v CR ERPs. Topographical maps of amplitude differences (top rows, blue/white/red colour map) and t-values for the differences (bottom rows, cold/black/hot colour map) between Hits and CRs split by Age Group and the final two rows showing the between groups difference. Only significant clusters ($p < .025$ at the cluster level) are shown in the t-statistic topographical maps.

When exploring how the Hits vs. CR difference interacted with age group (calculated based on a t-test comparing the Younger group Hits-CRs difference with the Older group Hits-CRs difference), three significant clusters emerged. The first was an early positive going cluster (*p*=.012) between approximately 550-750ms, indicating a larger Hits>CRs effect across left central electrodes in the Younger Adults than the Older Adults. Two later negative going clusters appeared to capture early and late parts of the same mid-frontal ERP effects. The first was a negative going mid-frontal cluster (*p*=.007) between around 1450-1600ms, with the second cluster similarly being widespread across the mid-frontal regions (*p*=.006) between around 1650-1850ms. This effect therefore captured the opposite going Hit-CR difference in the Younger vs. Older groups in the late part of the epoch (as described above, see Fig. 5.3).

Test 1 Foil ERP effects: FA v Fast CR

Next, I analysed Test 1 ERP differences between New foil items that were incorrectly identified as 'Old' (FAs) versus New items that were quickly correctly rejected (Fast CRs), to test the pre-registered predictions that FAs would be associated with an increased early FN400 positivity due to false/pre-experimental familiarity, whereas Fast CRs would be associated with an enhanced subsequent P3 response related to novelty detection, when compared to FAs. Grand-average ERPs from the four electrode sites for each condition are shown in Figure 5.4.

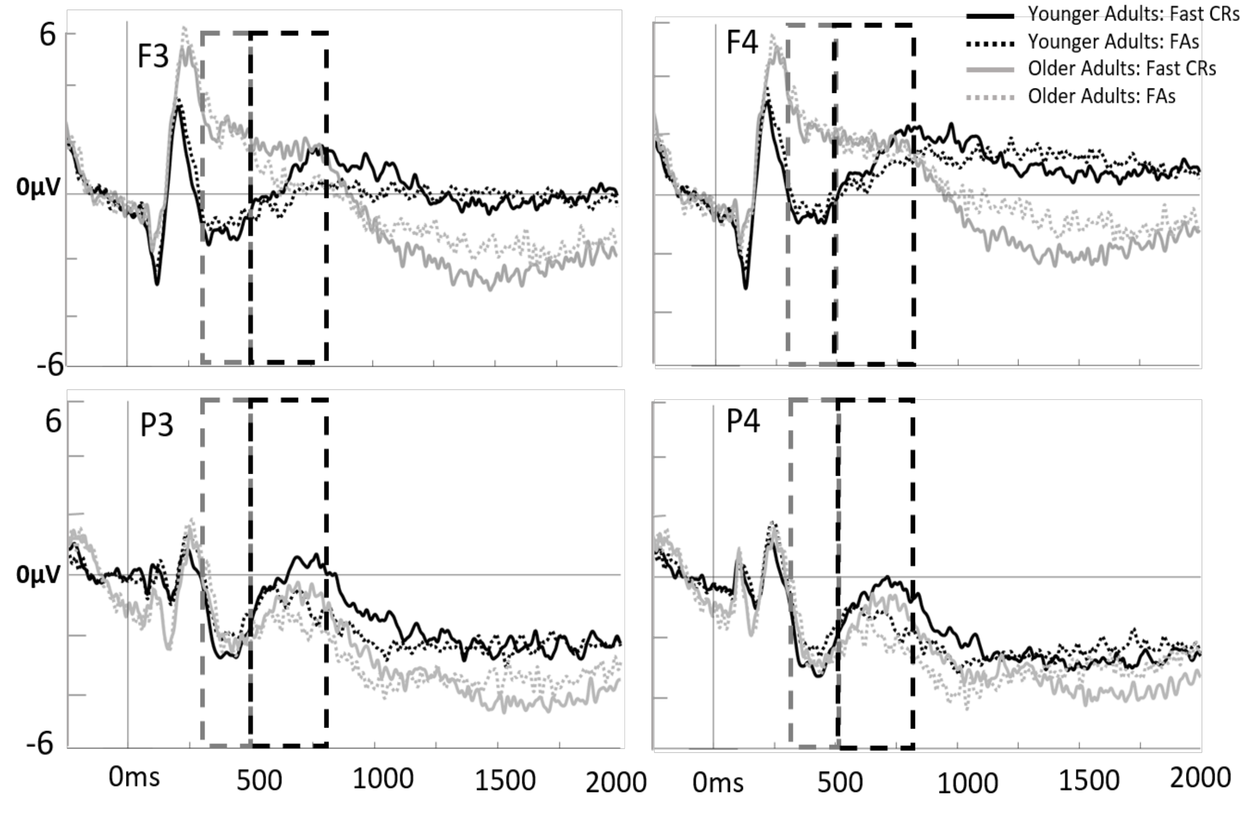


Figure 5.4. Grand-average ERPs for FAs and Fast CRs in Test 1 for Younger and Older Adults at left and right frontal (F3, F4) and left and right parietal (P3, P4) electrode sites. Boxes highlight the early 300-500ms and the late 500-800ms targeted time windows.

Results from the omnibus ANOVAs (factors; Condition: FA v Fast CR; Anterior/Posterior; Hemisphere and Age Group, see Table 5.15) in the early time window (300-500ms) showed a trend towards an enhanced positivity for FAs compared to Fast CRs, in line with the pre-registered directional effect. Since this effect would be significant with a one-tailed test (which can be argued to be justified with a directional, pre-registered hypothesis), it can be tentatively interpreted as a neural correlate of 'false' familiarity that may underlie erroneous 'Old' responses to these New items on Test 1. Interestingly, this effect had a broader/more posterior distribution than typical frontal FN400 effects and was similar in both Age Groups. In the later time-window (500-800ms), there was a significant 4-way interaction between FA vs. Fast CR Condition, Anterior/Posterior, Hemisphere and Age Group factors.

Table 5.15. *Results from a mixed ANOVA for the omnibus test comparing Age effects on ERPs associated with FAs vs. Fast CRs during Test 1 in the two a-priori targeted time windows.*

Omnibus	Time Window					
	300-500ms			500-800ms		
	F	*p*	n_p^2	*F*	*p*	n_p^2
C	3.07	0.085	0.05	**18.70**	**<.001**	**0.24**
C x AP	0.69	0.410	0.01	0.57	0.453	0.01
C x H	0.21	0.649	<.01	**5.55**	**0.022**	**0.09**
C x AP x H	0.21	0.650	<.01	**9.30**	**0.003**	**0.13**
Age x C	0.94	0.336	0.02	0.02	0.894	<.01
Age x C x AP	1.09	0.302	0.02	<001	1.000	<.01
Age x C x H	0.21	0.646	<.01	0.82	0.370	0.01
Age x C x AP x H	0.96	0.330	0.02	**8.76**	**0.004**	**0.13**

Note: C= Condition; AP =Anterior-Posterior; H = Hemisphere; Age = Age Group. Significant results are in bold. Only effects that involve the condition factor are reported. N=62

Follow up 3-way ANOVAs (C x AP x H) within each Age Group revealed differing effects of Condition within each Age Group (Table 5.16). Within the Younger Adults there was a significant main effect of Condition whereby Fast CRs were associated with more positive ERPs than FAs, but this effect did not differ depending on hemisphere or anteriority of the electrode locations.

Table 5.16. *Results from within groups ANOVAs, testing ERP differences between FAs and Fast CRs in the later time windows within Younger and Older groups separately.*

Condition Comparison	Age Group	Effect	Time Window					
			300-500ms			500-800ms		
			F	*p*	n_p^2	*F*	*p*	n_p^2
T1: FA v Fast CR	Younger Adults	C	-	-	-	**7.55**	**0.010**	**0.20**
		C x AP	-	-	-	0.27	0.606	0.01
		C x H	-	-	-	1.14	0.294	0.04
		C x AP x H	-	-	-	0.01	0.945	<.01
	Older Adults	C	-	-	-	**13.81**	**0.001**	**0.32**
		C x AP	-	-	-	0.31	0.585	0.01
		C x H	-	-	-	**4.91**	**0.035**	**0.14**
		C x AP x H	-	-	-	**15.23**	**0.001**	**0.34**

Note: C= Condition; AP =Anterior-Posterior; H = Hemisphere; Age = Age Group. Significant results are in bold. Only effects that involve the condition factor are reported. Younger Adults N=32, Older Adults N=30

However, within the older Adults there was a significant 3-way interaction which was followed-up with paired t-tests at each of the four electrode sites to directly investigate this interaction (Table 5.17).

Table 5.17. *Results from the follow up t-tests between FAs and Fast CRs at each electrode site for the late time-window for the older Adults only.*

Condition Comparison	Electrode	Time Window							
		300-500ms				500-800ms			
		t	*p*	*d*	BF_{10}	*t*	*p*	*d*	BF_{10}
T1: FA v Fast CR; within Older Adults	F3	-	-	-	-	**-4.54**	**<.001**	**-0.34**	**208.86**
	F4	-	-	-	-	-0.64	0.528	-0.06	1/4.42
	P3	-	-	-	-	**-3.46**	**0.002**	**-0.25**	**20.91**
	P4	-	-	-	-	**-3.32**	**0.002**	**-0.30**	**15.05**

Note: Significant results are in bold. N=30

The follow up t-tests showed an increased positivity for Fast CRs compared to FAs at the left front and both parietal electrode sites. This finding is therefore consistent with the prediction that items that are quickly and accurately responded to as ‘New’ engage an additional novelty-related process, reflected in the P3 effect; rather than items that are incorrectly identified as ‘Old’. This process appears to not severely affected by ageing since it was similar across both Age Groups, but with the older adults showing a slightly different topography.

Consistent with the targeted analysis, the cluster-based permutation analysis (Fig. 5.5) indicated a highly significant ERP differences between FAs and Fast CRs within the Younger Adults with ERPs for Fast CRs more positive than FAs (*p*=.002) between approximately 650ms-870ms across central and parietal electrode sites. Within the Older Adults there was a very similar cluster (*p*=.002) approximately around 650ms-800ms, and there was no interaction with Age group for this (or any other) ERP effect. These effects are consistent with a novelty-related P3 for Fast CRs as predicted, and suggest that similar novelty processes occur in both age groups.

Within the Older Adults, two further clusters indicated that ERPs for FAs were more positive than Fast CRs in the later part of the epoch. Both of these late clusters were positive and broadly distributed, with the first (*p*=.011) extending approximately between 1440ms-1500ms and the second (*p*=.002) lasting for longer approximately between 1520ms-1640ms (as seen in Fig 5.5), and thus they appeared to capture early and later parts of the same ERP effect. This late increased positivity for FAs compared to Fast CRs did not interact with Age Group, since Younger Adults showed a numerical difference in the same direction, but this was smaller and non-significant.

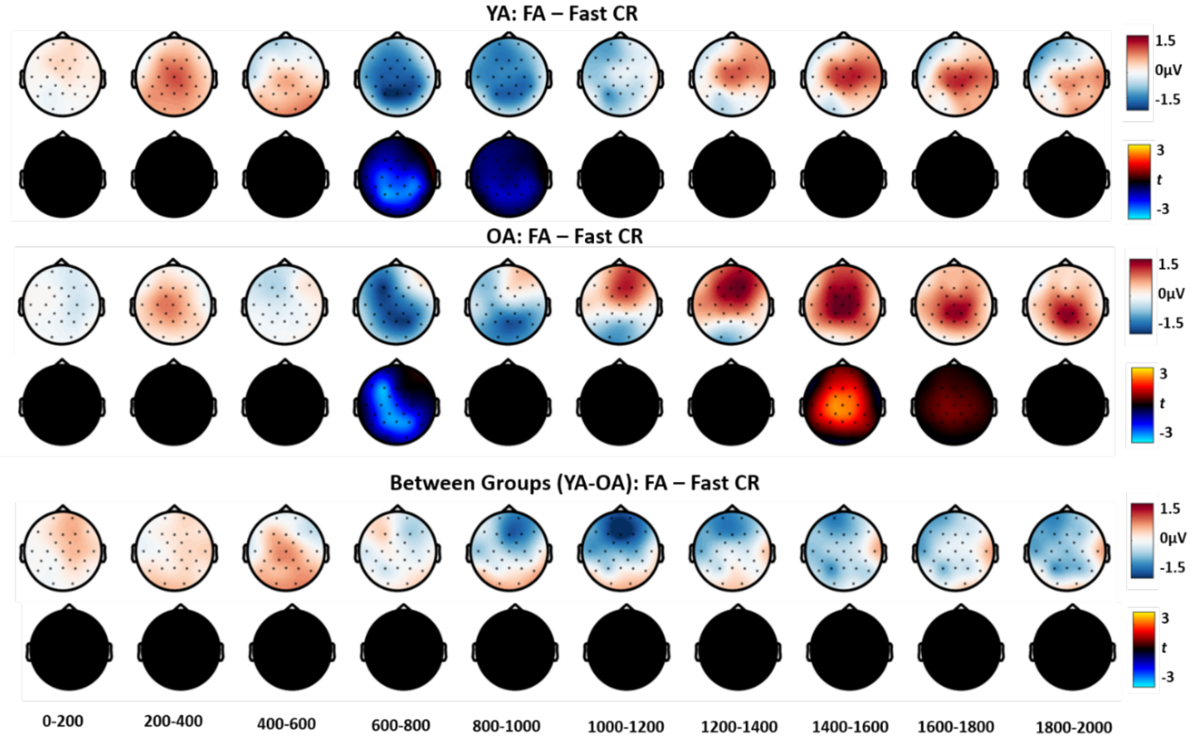


Figure 5.5. Whole-scalp analysis of Test 1 FAs v Fast CR ERPs. Topographical maps of amplitude differences (top rows, blue/white/red colour map) and t-values for the differences (bottom rows, cold/black/hot colour map) between FAs v Fast CR split by Age Group and the final two rows showing the between groups difference. Only significant clusters (p < .025 at the cluster level) are shown in the t-statistic topographical maps.

Test 1 Foil ERP effects: FA v Slow CR

Next, I analysed Test 1 ERP differences between New foil items that were incorrectly identified as 'Old' (FAs) versus New items that were slowly correctly rejected (Slow CRs), to test the pre-registered predictions that FAs would be associated with an increased early FN400 positivity compared to both CR conditions due to false/pre-experimental familiarity. For Slow CRs, I did not expect to see an enhanced P3 effect as found for Fast CRs, if that P3 effect is related to rapid novelty detection as hypothesised. Grand-average ERPs from the four electrode sites for each condition are shown in Figure 5.6.

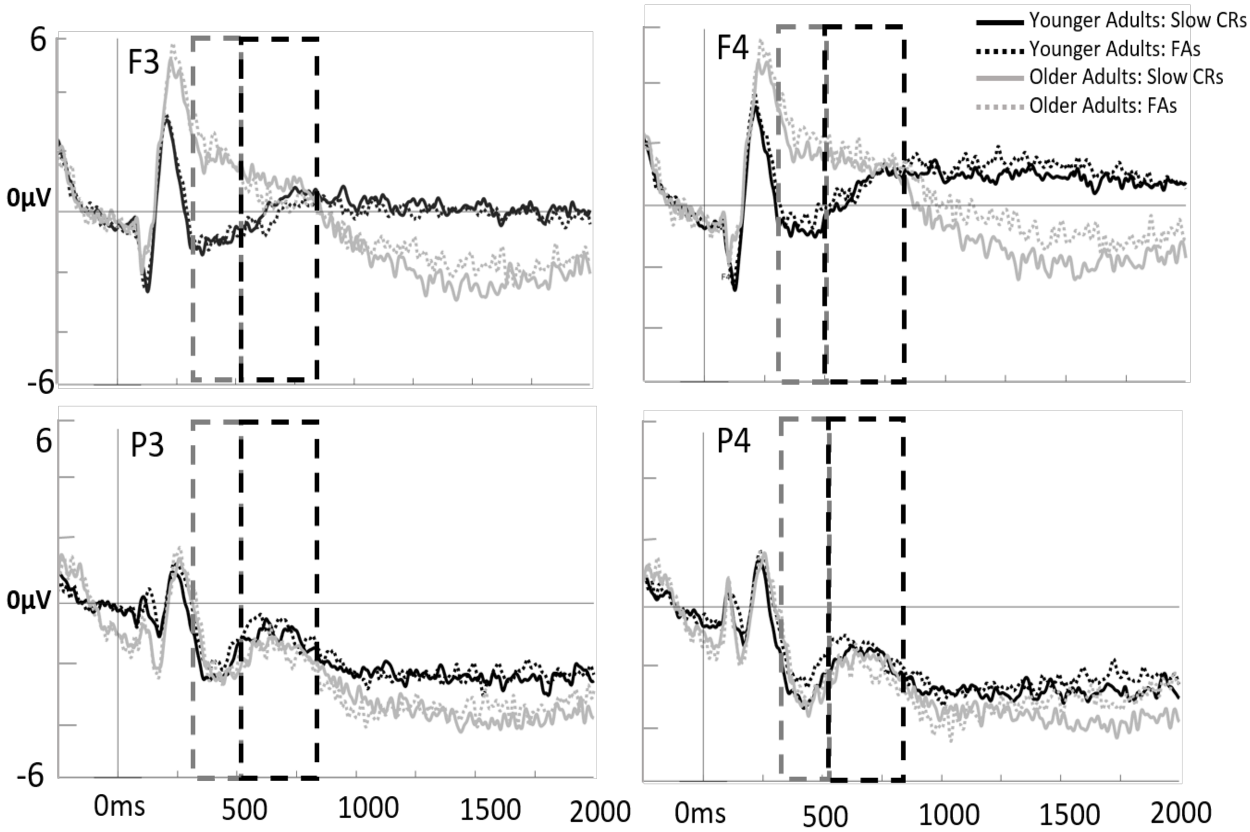


Figure 5.6. Grand-average ERPs from FAs and Slow CRs from Test 1 for Younger and Older Adults at F3, F4, P3, P4. Boxes highlight the early 300-500ms and the late 500-800ms targeted time windows.

Results from the omnibus ANOVAs (factors; Condition: FA v Slow CR; Anterior/Posterior; Hemisphere and Age Group) showed a significant main effect of Condition (Table 5.18) in the early time-window, that did not interact with Age or electrode location factors. ERPs for FAs were more positive than ERPs for Fast CRs, in line with the pre-registered directional effect, based on the hypothesis that 'false' familiarity as reflected in the FN400 may underlie erroneous 'Old' responses on Test 1.

Table 5.18. *Results from a mixed ANOVA for the omnibus test comparing Age effects on ERPs associated with FAs vs. Slow CRs during Test 1 in the two a-priori targeted time windows.*

Omnibus	Time Window					
	300-500ms			500-800ms		
	F	*p*	n_p^2	*F*	*p*	n_p^2
C	**7.09**	**0.010**	**0.11**	0.05	0.817	<.01
C x AP	0.02	0.898	<.01	0.68	0.413	0.01
C x H	3.31	0.074	0.05	**8.50**	**0.005**	**0.12**
C x AP x H	1.16	0.287	0.02	**14.01**	**<.001**	**0.19**
Age x C	0.15	0.703	<.01	0.30	0.586	0.01
Age x C x AP	2.87	0.096	0.05	0.60	0.441	0.01
Age x C x H	0.29	0.595	<.01	0.20	0.657	<.01
Age x C x AP x H	0.04	0.839	<.01	**4.05**	**0.049**	**0.06**

Note: C= Condition; AP =Anterior-Posterior; H = Hemisphere; Age = Age Group. Significant results are in bold. Only effects that involve the condition factor are reported. N=62

The omnibus ANOVA also showed a significant 4-way interaction in the late time-window (Table 5.18) which was followed up with 3-way ANOVAs within each Age Group (Table 5.19).

Table 5.19. *Results from within groups ANOVAs, testing ERP differences between FAs and Slow CRs in the later time windows within Younger and Older groups separately*

Condition Comparison	Age Group	Effect	Time Window					
			300-500ms			500-800ms		
			F	*p*	n_p^2	*F*	*p*	n_p^2
T1: FA v Slow CRs	Younger Adults	C	-	-	-	0.03	0.853	<.01
		C x AP	-	-	-	1.02	0.320	0.03
		C x H	-	-	-	2.71	0.110	0.08
		C x AP x H	-	-	-	1.52	0.227	0.05
	Older Adults	C	-	-	-	0.59	0.447	0.02
		C x AP	-	-	-	<.01	0.968	<.01
		C x H	-	-	-	**6.64**	**0.015**	**0.19**
		C x AP x H	-	-	-	**16.36**	**<.001**	**0.36**

Note: C= Condition; AP =Anterior-Posterior; H = Hemisphere; Age = Age Group. Significant results are in bold. Only effects that involve the condition factor are reported. Younger Adults N=32, Older Adults N=30

There were no differences between FAs and Slow CRs in ERP amplitudes between 500-800ms in the Younger Adults, indicating that the 4-way interaction was driven by effects in the Older group. In the Older group, there was a further 3-way interaction between Condition (FA v Slow CRs), Anterior/Posterior and Hemisphere which was followed up with paired t-tests conducted at each electrode site (Table 5.20). These showed more positive left frontal ERPs for Slow CRs than FAs within the older group.

Table 5.20. *Results from the follow up t-tests between FAs and Slow CRs at each electrode site for the late time-window within the Older Adult group.*

Condition Comparison	electrode	Time Window							
		300-500ms				500-800ms			
		t	*p*	*d*	BF_{10}	*t*	*p*	*d*	BF_{10}
T1: FA v Slow CR; Older Adults	F3	-	-	-	-	**-2.11**	**0.043**	**-0.18**	**1.34**
	F4	-	-	-	-	0.90	0.377	0.08	1/3.56
	P3	-	-	-	-	-0.35	0.728	-0.03	1/4.86
	P4	-	-	-	-	-0.89	0.380	-0.07	1/3.58

Note: Significant results are in bold. N=30

The early FN400 like effects found in the targeted time-window analysis did not survive the correction for multiple comparisons in the cluster-based permutation analysis. The only significant clusters were within the Older Adults, where there was a significant late positive cluster (p<.001) approximately between 1520ms-1640ms (as seen in Fig 5.7), where FAs were associated with more positive ERPs than Slow CRs (similarly to the late positive effect found when comparing FAs and Fast CRs). However, this effect did not interact with age group, and there were no significant effects within the younger group.

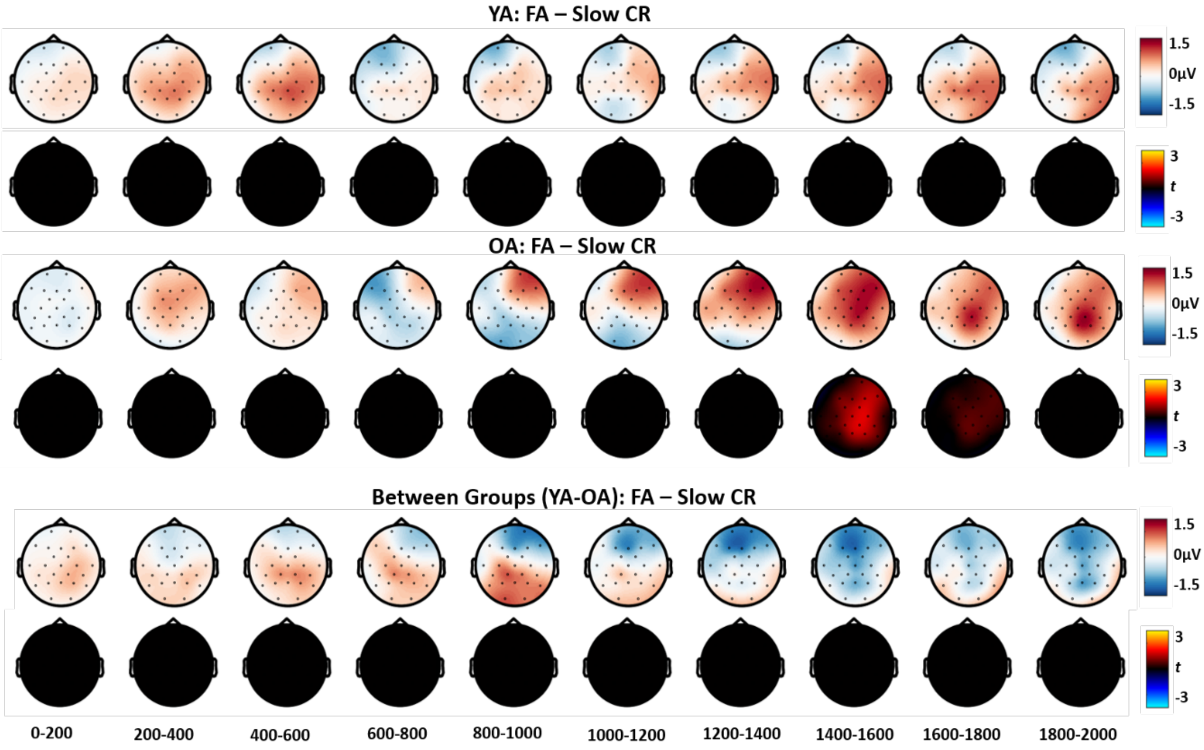


Figure 5.7. Whole-head analysis of Test 1 FAs v Slow CR ERPs. Topographical maps of amplitude differences (top rows, blue/white/red colour map) and t-values for the differences (bottom rows, cold/black/hot colour map) between FAs v Slow CR split by Age Group and the final two rows showing the between groups difference. Only significant clusters (p < .025 at the cluster level) are shown in the t-statistic topographical maps.

Test 1 Foil ERP effects: Fast CR v Slow CR

In the final analysis of Test 1 ERPs, I compared Fast and Slow CRs directly to test for differences in neurocognitive processes between these types of CRs. I expected a larger P3 effect for Fast CRs compared to Slow CRs, if the P3 effect is related to rapid novelty detection. Grand-average ERPs from the four electrode sites for each condition are shown in Figure 5.8.

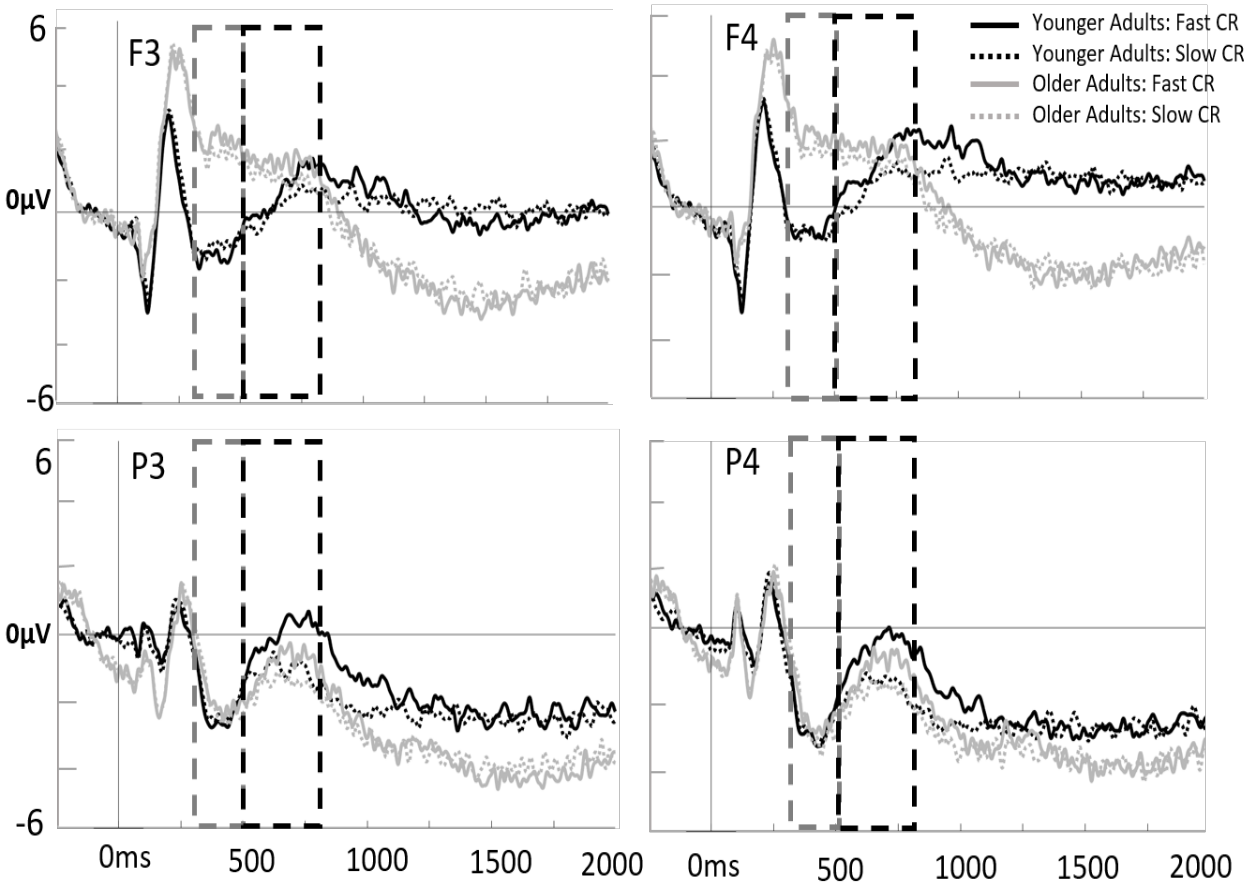


Figure 5.8. Grand-average ERPs from Fast CRs and Slow CRs from Test 1 for Younger and Older Adults at F3, F4, P3, P4. Boxes highlight the early 300-500ms and the late 500-800ms targeted time windows.

Results from the omnibus ANOVA (factors; Condition: Fast CR v Slow CR; Anterior/Posterior; Hemisphere and Age Group) are shown in Table 5.21. In the early time window (300-500ms) were no significant ERP differences, consistent with the prediction that there should not be any large differences in early familiarity effects between two types of CRs, since enhanced familiarity would primarily lead to FA responses.

In the later time window (500-800ms), there was a significant main effect of Condition with Fast CRs being more positive than Slow CRs overall, and this effect did not differ between age groups or across electrode locations. This effect is thus consistent with an enhanced P3 for Fast compared to Slow CRs across both younger and older adults.

Table 5.21. *Results from a mixed ANOVA for the omnibus test comparing Age effects on ERPs associated with Fast CRs vs. Slow CRs during Test 1 in the two a-priori targeted time windows.*

Omnibus	Time Window					
	300-500ms			500-800ms		
	F	*p*	n_p^2	*F*	*p*	n_p^2
C	0.27	0.607	<.01	**17.20**	**<.001**	**0.22**
C x AP	0.36	0.553	0.01	2.94	0.091	0.05
C x H	2.94	0.092	0.05	0.24	0.627	<.01
C x AP x H	0.41	0.525	0.01	0.80	0.374	0.01
Age x C	0.68	0.412	0.01	0.57	0.454	0.01
Age x C x AP	0.48	0.491	0.01	0.78	0.380	0.01
Age x C x H	1.22	0.274	0.02	0.27	0.607	<.01
Age x C x AP x H	0.90	0.348	0.02	2.74	0.103	0.04

Note: C= Condition; AP =Anterior-Posterior; H = Hemisphere; Age = Age Group. Significant results are in bold. Only effects that involve the condition factor are reported. N=62

The cluster-based permutation analysis (Fig. 5.9) indicated highly significant ERP differences between Fast CRs and Slow CRs only in the Younger Adult group, with a significant positive cluster (p<.001) approximately around 640ms-1020ms due to ERPs for Fast CRs being more positive than ERPs for Slow CRs, consistent with the P3 effect in the targeted analysis. This effect was lower in amplitude and not significant in the Older Adults, but there was no significant interaction with Age group.

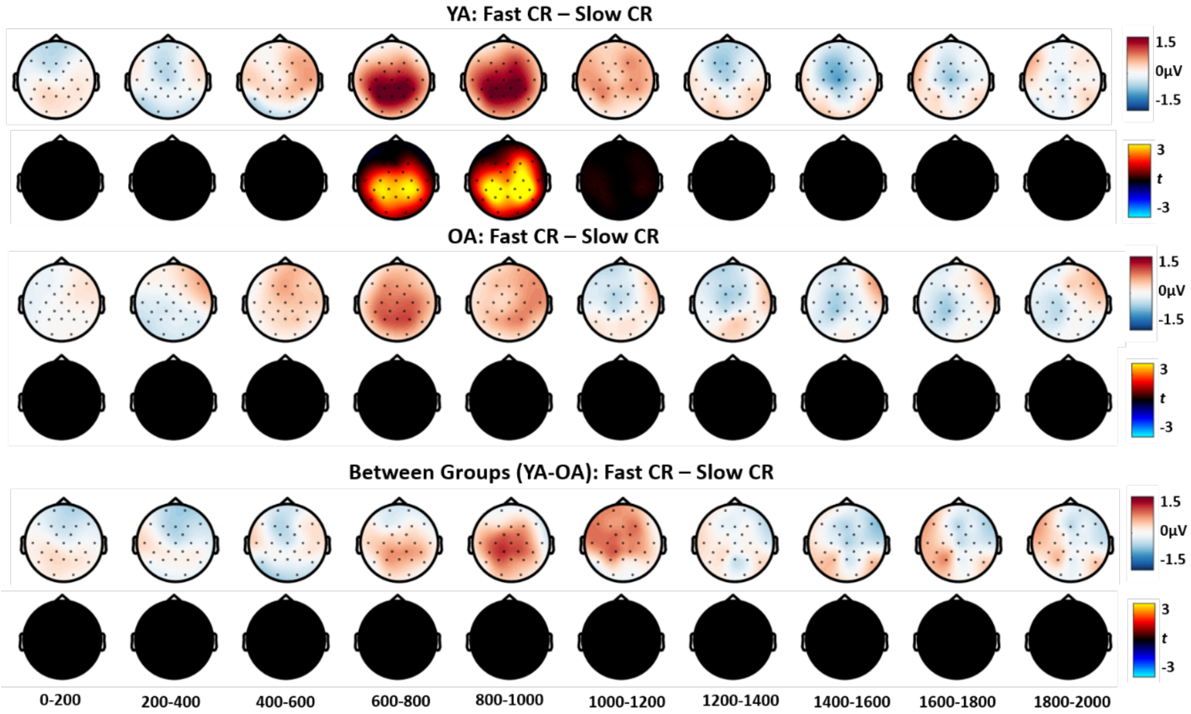


Figure 5.9. Whole-head analysis of Test 1 Fast CR v Slow CR ERPs. Topographical maps of amplitude differences (top rows, blue/white/red colour map) and t-values for the differences (bottom rows, cold/black/hot colour map) between Fast CR v Slow CR split by Age Group and the final two rows showing the between groups difference. Only significant clusters (p < .025 at the cluster level) are shown in the t-statistic topographical maps.

Recognition Test 1 Summary of ERP effects

ERPs during recognition test 1 showed several predicted ERP effects for successful recognition (hits compared to correct rejections), with ERP components associated with familiarity and recollection for old hits. There were some age differences in these effects however, with older adults showing a similar, but reduced and more frontal pattern of activity, whereas younger adults showed the standard widespread increased positivity for hits compared to correct rejections. Whole-head cluster analysis revealed group differences in activity in the later part of the epoch, with a reversal in polarity. Younger adults had a negative going slow-wave for old hits with new CRs being more positive, whereas older adults showed a positive going slow-wave with increased positivity for hits.

In relation to the hypothesised ERP differences associated with recognition decisions for new items, early effects in the FN400 time-window were seen for FAs compared to Fast and Slow CRs (although some of these differences were marginally significant) that appeared similar across age groups. As predicted, both age groups showed later effects in the P3 time-window for Fast CRs compared to FAs, with the older adults having a more left-frontal and parietal focussed ERP effect. Interestingly, going against predictions, there was an increased effect in the P3 time-window for Slow CRs compared to FAs only within the Older adults. However no between-group differences were seen in this effect, so it cannot be interpreted as age-specific. When comparing Fast CRs to Slow CRs there was an increased positivity for Fast CRs in line with the prediction that these will induce a larger effect in the P3 time-window. No between group differences in this comparison were seen, except when examined with the cluster analysis this effect was not present for Older Adults. However again no between-group differences in the Fast vs. Slow CR difference, so this effect cannot be interpreted as age-sensitive.

Finally, whole-scalp cluster analysis revealed a significant cluster in the Older adults late in the epoch with increased positivity for FAs compared to both Fast CRs and Slow CRs. Younger adults showed a similar directional effect but this was not significant, and there was no significant age difference in this effect.

5.4.3.2 Surprise Recognition Test ERP Results

The analysis in this section examined ERPs related to the recognition of foil words that were encoded during Recognition Test 1.

Test 2 Correct Recognition ERP effects: Surprise Test Hits vs. Surprise Test CRs

In the first analysis of the surprise test ERPs, I compared ERPs for successfully recognised prior foils (which in this test were old items, here referred to as ST Hits) with ERPs for successfully rejected new items (New CRs). This analysis thus predicts similar old/new ERP effects as in the comparison between Hits and CRs on Test 1, with the difference that the Test 1 comparison showed recognition effects after intentional encoding, whereas in the surprise test the ‘Old’ Foil items were incidentally encoded. Nevertheless, I expected to see an early (~300-500ms) familiarity-related FN400 modulation for ST Hits compared to New CRs, with decreased negativity for ST Hits. I would also expect to see a parietal Old/New ERP positivity for ST Hits compared to New CRs in the 500-800ms time-window. Since behavioural performance differed between the two age groups, it was of interest to test whether that would result in differences in FN400 and parietal old/new ERP effects between age groups. Grand-average ERPs from the four electrode sites are shown in Figure 5.10.

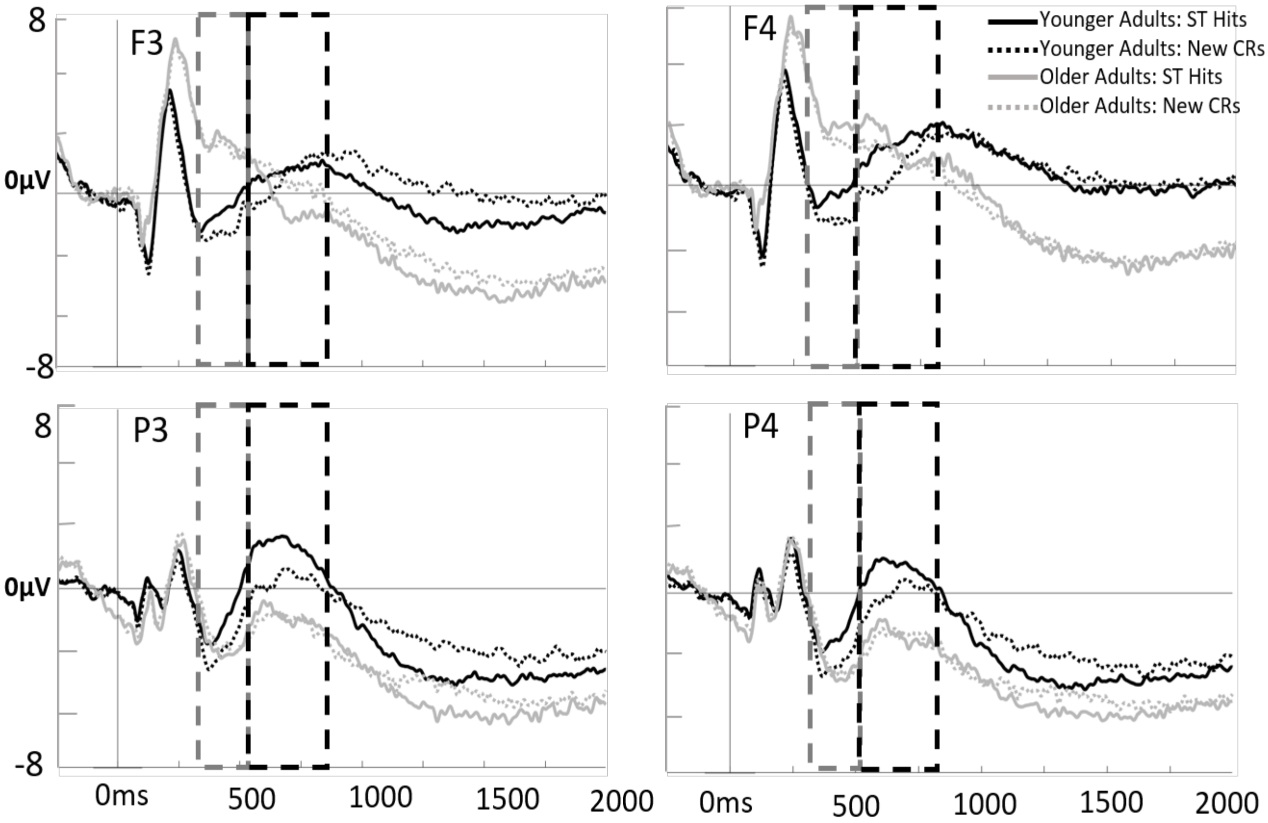


Figure 5.10. Grand-average ERPs from ST Hits and New CRs from the Surprise Test for Younger and Older Adults at F3, F4, P3, P4. Boxes highlight the early 300-500ms and the late 500-800ms targeted time windows.

Results from the omnibus ANOVAs (factors; Condition: ST Hits v New CRs; Anterior/Posterior; Hemisphere and Age Group) are shown in Table 5. In the early time-window (300-500ms) there were two significant 3-way interactions, one between Condition (ST Hits v New CR), Anterior/Posterior and Hemisphere and a second between Condition (ST Hits v New CR), Anterior/Posterior and Age Group. In the later time window (500-800ms) there were similarly two significant 3-way interactions; one between Condition (ST Hits v New CR), Anterior/Posterior and Hemisphere and a second between Condition (ST Hits v New CR), Hemisphere and Age Group.

Table 5.22. *Results from a mixed ANOVA for the omnibus test comparing Age effects on ERPs associated with ST Hits vs New CRs during the Surprise Test in the two a-priori targeted time windows.*

Omnibus	Time Window					
	300-500ms			500-800ms		
	F	*p*	n_p^2	*F*	*p*	n_p^2
C	**14.61**	**<.001**	**0.20**	**9.36**	**0.003**	**0.13**
C x AP	3.59	0.063	0.06	**7.68**	**0.007**	**0.11**
C x H	**7.10**	**0.010**	**0.11**	**12.58**	**0.001**	**0.17**
C x AP x H	**4.27**	**0.043**	**0.07**	**75.59**	**<.001**	**0.56**
Age x C	**13.35**	**0.001**	**0.18**	**11.77**	**0.001**	**0.16**
Age x C x AP	**9.43**	**0.003**	**0.14**	0.46	0.500	0.01
Age x C x H	0.69	0.409	0.01	**6.46**	**0.014**	**0.10**
Age x C x AP x H	0.21	0.650	<.01	2.56	0.115	0.04

Note: C= Condition; AP =Anterior-Posterior; H = Hemisphere; Age = Age Group. Significant results are in bold. Only effects that involve the condition factor are reported. N=62

To follow up on the 3-way interaction between Condition (ST Hits v New CR), Anterior/Posterior and Hemisphere observed in both time-windows paired t-tests were conducted at each electrode site across all participants (Table 5.23).

Table 5.23. *Results from the follow up t-tests at each electrode site for the early time-window (300-500ms) and the late time-window for all participants.*

Condition Comparison	electrode	Time Window							
		300-500ms				500-800ms			
		t	*p*	*d*	BF_{10}	*t*	*p*	*d*	BF_{10}
ST Hit v New CR	F3	**2.52**	**0.014**	**0.11**	**2.52**	-0.88	0.381	-0.06	1/4.96
	F4	**4.71**	**<.001**	**0.20**	**1243.86**	**4.03**	**<.001**	**0.25**	**141.78**
	P3	**2.22**	**0.031**	**0.12**	**1.34**	**4.37**	**<.001**	**0.27**	**410.77**
	P4	**3.30**	**0.002**	**0.16**	**17.35**	**3.08**	**0.003**	**0.21**	**9.59**

Note: Significant results are in bold. N=62

These showed that in the early time-window (300-500ms) there were widespread significant effects with ST Hits being more positive than New CRs, but the difference was maximal over the right frontal electrode. In the later time-window (500-800ms) there was similar widespread positivity for ST Hits compared to New CRs except for in the frontal left region where no significant differences were observed, and this later effect was now strongest at the left parietal electrode.

To follow up on the 3-way interaction between Condition (ST Hits v New CR), Anterior/Posterior and Age Group in the early time-window (300-500ms) paired t-tests were conducted in the frontal and parietal regions separately for each Age Group, collapsed across the hemisphere factor. Within the Younger Adults ST Hits were more positive than New CRs in both the frontal ($t(31)=3.51$, $p=.001$, $d=.28$, $BF_{10}=24.25$) and parietal regions ($t(31)=5.74$, $p<.001$, $d=.35$, $BF_{10}=7260.70$) with the effect being maximal over the parietal regions. In the Older Adults, there was no significant difference between conditions in the frontal region ($t(29)=1.57$, $p=.126$, $d=.09$, $BF_{10}=1/1.71$), and only a trend level difference in the parietal region ($t(29)=1.90$, $p=.068$, $d=.11$, $BF_{10}=1/1.06$) showing a reversed tendency towards increased positivity for New CRs over ST Hits.

To follow up on the 3-way interaction between Condition (ST Hits v New CR), Hemisphere and Age Group in the late time-window (500-800ms) paired t-tests were conducted in the left and right hemispheres separately for each Age Group, collapsed across the Anterior-Posterior factor. Within the Younger Adults there were significantly increased ERP positivity for ST Hits compared to New CRs in both the left ($t(31)=3.87$, $p=.001$, $d=.37$, $BF_{10}=57.56$) and right hemispheres ($t(31)=4,.11$, $p<.001$, $d=.46$, $BF_{10}=104.26$).Older Adults however did not show significant differences in either Hemisphere (Left: $t(29)=1.74$, $p=.093$, $d=.15$, $BF_{10}=1/1.34$; Right: $t(29)=1.25$, $p=.221$, $d=.11$, $BF_{10}=1/2.53$). Thus, the follow up analyses suggested that only the younger group showed a consistent enhanced positivity for ST Hits compared to New CRs in the 300-500ms and 500-800ms time-windows.

Consistently with the targeted analysis, the non-parametric cluster-based permutation analysis showed a significant age group difference in the ST Hit vs. New CRs effect (P<.001) around 320-790ms (as seen in Fig 5.11). Whereas the Younger Adults had a highly significant broadly distributed cluster (p<.001) with more positive ERPs for ST Hits compared to New CRs approximately around 340ms-790ms, the Older Adults however showed no significant differences between conditions around this time. However, there were similarities across the Age Groups in the later stages of the epoch with both groups showing increased ERP positivity for New CRs compared to ST Hits. The Younger Adults had a negative cluster (p<.001) approximately around 1040-1730ms, with the Older Adults also having a negative cluster (p=.002) around 1160-1500ms, which had similar topographies in the two groups.

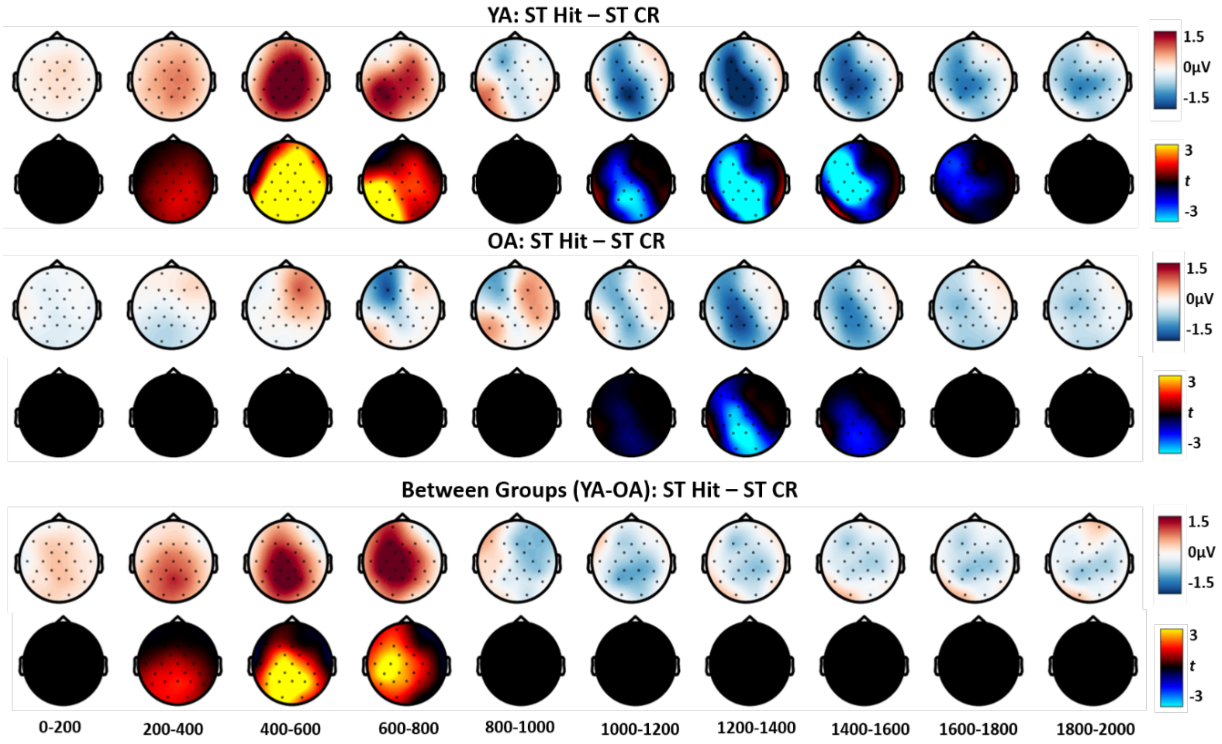


Figure 5.11. Whole-head analysis of ST Hits v New CR ERPs. Topographical maps of amplitude differences (top rows, blue/white/red colour map) and t-values for the differences (bottom rows, cold/black/hot colour map) between ST Hits v New CR split by Age Group and the final two rows showing the between groups difference. Only significant clusters (p<.025 at the cluster level) are shown in the t-statistic topographical maps.

ST Hit given that foils were a Prior FA v Prior Fast CR

Next, I compared ERPs for Foil items that were correctly recognised on the surprise test (ST Hits) but previously received either a False Alarm (Prior FA) or a Fast CR (Prior Fast CR) response on the first test. The purpose of this analysis was to test if the neural processes involved in the initial recognition decision, such as a 'false' familiarity process, are reinstated on this subsequent test for the same items. In addition, it was hypothesised that retrieval- and decision-related processes engaged on the first test (such as novelty detection) could enhance encoding of items, which might therefore lead to differences in retrieval related ERPs on the final test (such as enhanced recollection-related effects for Prior Fast CRs as a result of novelty detection on the first test). Grand-average ERPs from the four analysed electrode sites for each condition are shown in Figure 5.12.

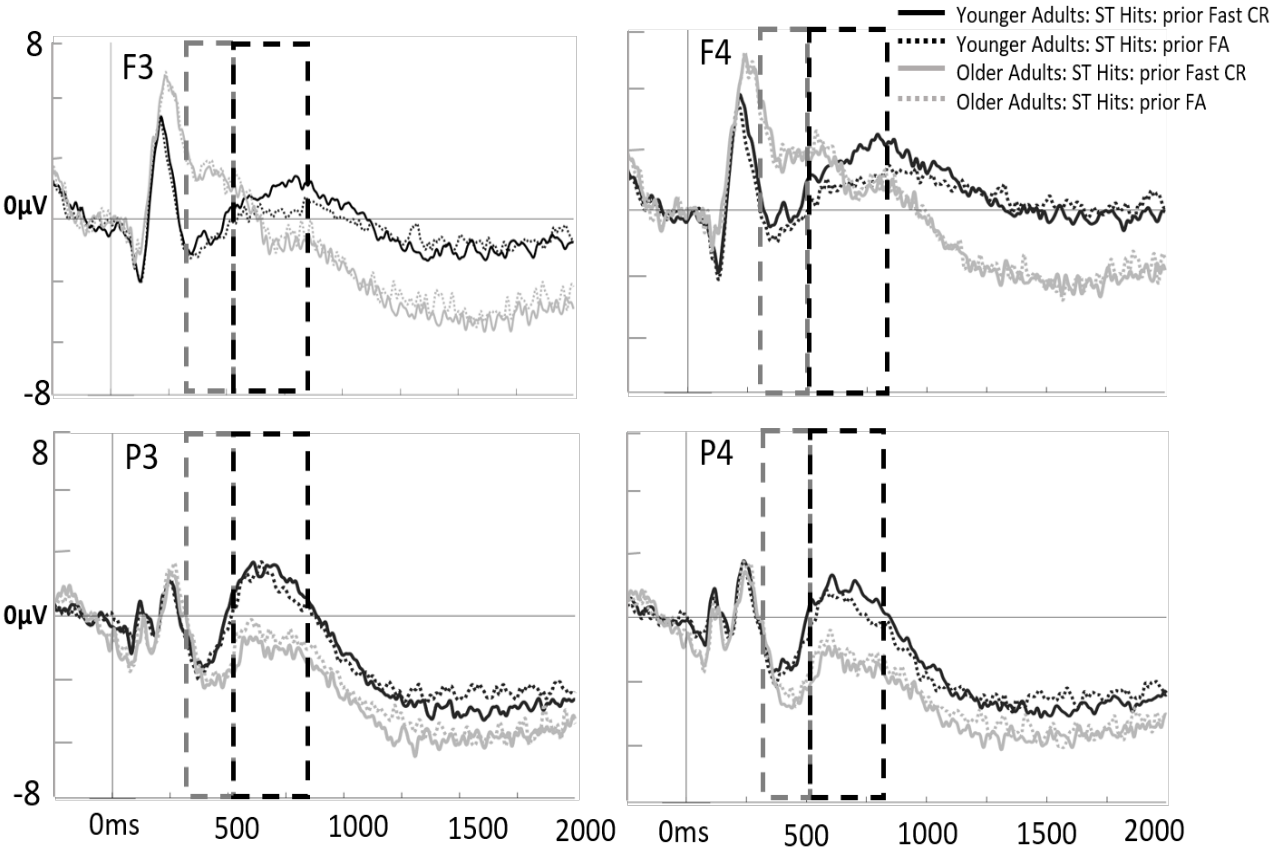


Figure 5.12. Grand-average ERPs for correctly recognised Prior FAs and Prior Fast CRs on the Surprise Test for Younger and Older Adults at F3, F4, P3, P4. Boxes highlight the early 300-500ms and the late 500-800ms targeted time windows.

Results from the omnibus ANOVAs (factors: Condition: Prior FA v Prior Fast CR; Anterior/Posterior; Hemisphere and Age Group) are shown in Table 5.24. In the early time window (300-500ms) there were no effects of Condition on ERP amplitudes, and no interactions between Condition and Age Group. In the later time window (500-800ms) there were two effects of interest however; Condition interacted with Hemisphere (C x H) and Condition interacted with Age Group (C x Age).

Table 5.24. *Results from a mixed ANOVA for the omnibus test comparing Age effects on ERPs associated with ST Hits: prior FA vs ST Hits: prior Fast CR during the Surprise Test in the two a-priori targeted time windows.*

Omnibus	Time Window					
	300-500ms			500-800ms		
	F	*p*	n_p^2	*F*	*p*	n_p^2
C	0.52	0.475	0.01	0.65	0.423	0.01
C x AP	1.60	0.211	0.03	2.42	0.125	0.04
C x H	3.40	0.070	0.05	**5.44**	**0.023**	**0.08**
C x AP x H	1.47	0.231	0.02	0.05	0.833	<.01
Age x C	1.55	0.219	0.03	**7.82**	**0.007**	**0.12**
Age x C x AP	0.03	0.872	<.01	0.36	0.554	0.01
Age x C x H	0.54	0.467	0.01	1.02	0.316	0.02
Age x C x AP x H	1.66	0.203	0.03	0.83	0.367	0.01

Note: C= Condition; AP =Anterior-Posterior; H = Hemisphere; Age = Age Group. Significant results are in bold. Only effects that involve the condition factor are reported. N=62

To follow up on the interaction between Condition and Hemisphere, paired t-tests were conducted to compare Prior FAs with Prior Fast CRs in Right (R) and Left (L) Hemispheres, collapsed across Anterior-Posterior and grouping factors. However, there were no significant differences between conditions in either hemisphere (Left: *t*(61)=0.44, *p*=.660, *d*=.03, BF_{10}=1/6.55; Right: *t*(61)=1.23, *p*=.225, *d*=.11, BF_{10}=1/3.53). To follow up on the interaction between Condition and Age Group, paired t-tests were conducted to compare Prior FAs with Prior Fast CRs collapsed across Anterior-Posterior and Hemisphere factors, separately for Younger Adults and Older Adults. Within the Younger Adults, Prior Fast CRs were more positive than Prior FAs (*t*(31)=2.33, *p*=.027, *d*=.28, BF_{10}=1.94). Within the Older Adults there was no significant amplitude differences between conditions (*t*(29)=1.61, *p*=.119, *d*=.17, BF_{10}=1/1.63).

Consistent with the targeted analysis, the global cluster-based permutation analysis indicated significantly more negative ERPs for Prior FAs than Prior Fast CRs in the Younger Adults (p<.001) approximately 690-870ms and maximal over frontal and central electrode sites. No significant clusters were observed in the Older Adults, however the interaction between groups was not significant but a trend level cluster was observed (*p*=.026 meaning it just failed to exceed the α-level of *p*<.025) approximately 730-810ms (as seen in Fig. 5.13).

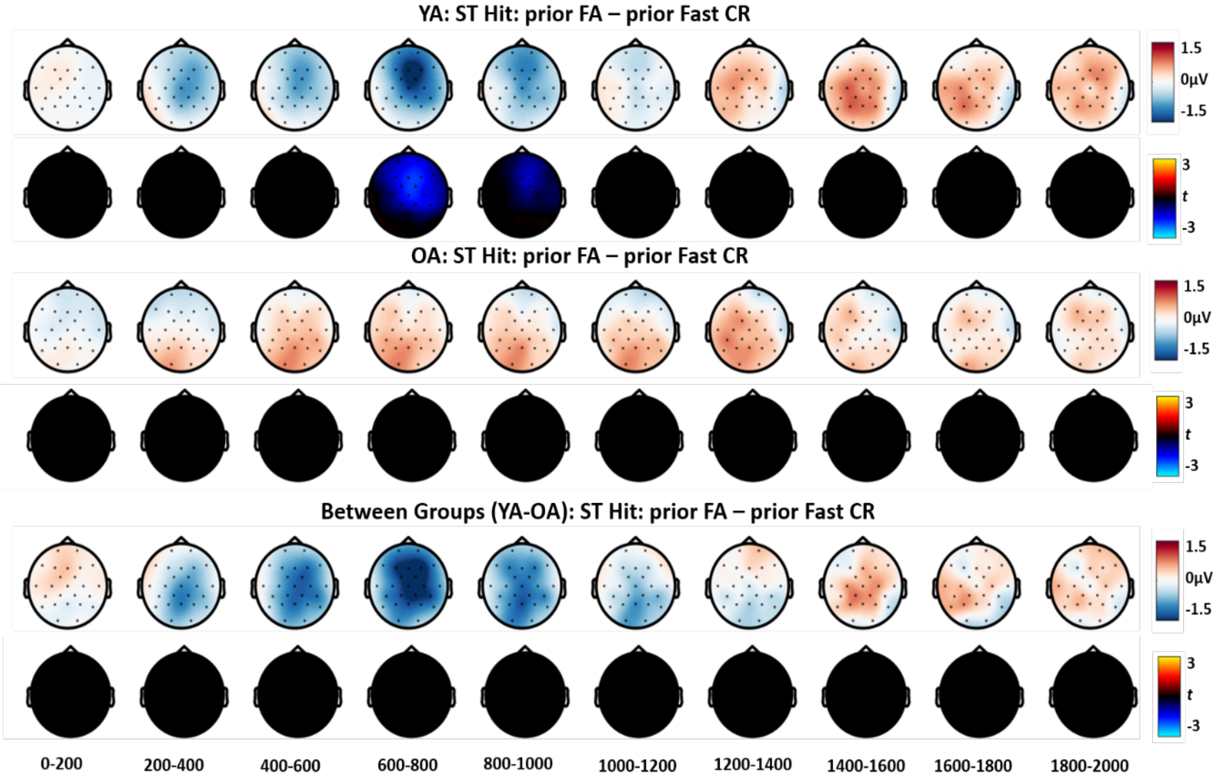


Figure 5.13. Whole-head analysis of ST Hits: prior FA v prior Fast CR ERPs. Topographical maps of amplitude differences (top rows, blue/white/red colour map) and t-values for the differences (bottom rows, cold/black/hot colour map) between ST Hits: prior FA v prior Fast CR split by Age Group and the final two rows showing the between groups difference. Only significant clusters (p < .025 at the cluster level) are shown in the t-statistic topographical maps.

ST Hit given that foils were a Prior FA v Prior Slow CR

Next, I compared ERPs for Foil items that were correctly recognised on the surprise test (ST Hits) but previously received either a False Alarm (Prior FA) or a Slow CR (Prior Slow CR) response on the first test. The purpose of this analysis was also to investigate if false familiarity and novelty processing on Test 1 resulted in different retrieval processing on Test 2. Grand-average ERPs from the four electrode sites for each condition are shown in Figure 5.14.

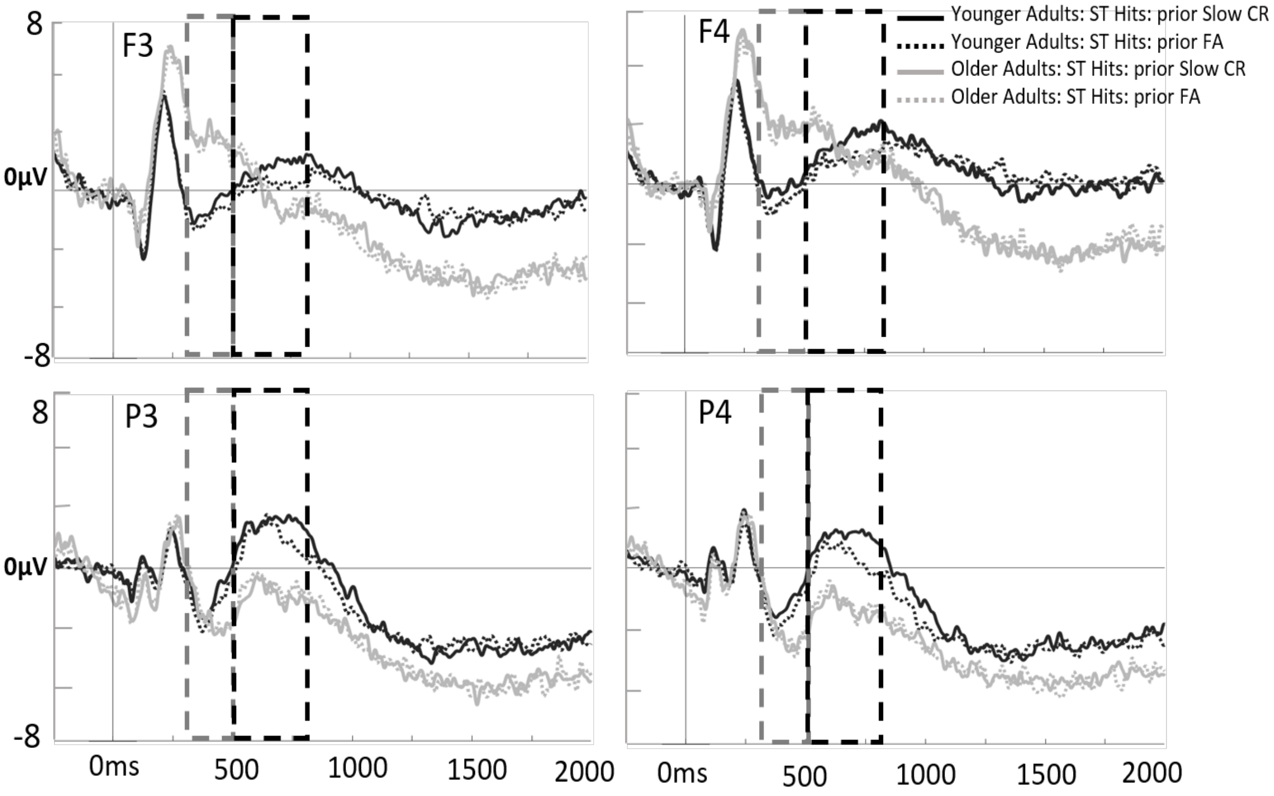


Figure 5.14. Grand-average ERPs from ST Hits; prior FA and ST Hits; prior Slow CR from the Surprise Test for Younger and Older Adults at F3, F4, P3, P4. Boxes highlight the early 300-500ms and the late 500-800ms targeted time windows.

Results from the omnibus ANOVAs (factors; Condition: Prior FA v Prior Slow CR; Anterior/Posterior; Hemisphere and Age Group) are shown in Table 5.25. In the early time window (300-500ms) there were no significant condition differences in amplitude, with only a trend level increase in positivity for prior Slow CRs compared to prior FAs. In the later time-window (500-800ms), a trend level interaction was seen between Condition and Age Group, similarly to in the analysis comparing Prior FAs with Prior Fast CRs.

Table 5.25. *Results from a mixed ANOVA for the omnibus test comparing Age effects on ERPs associated with ST Hits: prior FA vs ST Hits: prior Slow CR during the Surprise Test in the two a-priori targeted time windows.*

Omnibus	Time Window					
	300-500ms			500-800ms		
	F	*p*	n_p^2	*F*	*p*	n_p^2
C	3.01	0.088	0.05	1.13	0.293	0.02
C x AP	0.12	0.728	<.01	0.08	0.780	<.01
C x H	0.86	0.358	0.01	2.27	0.137	0.04
C x AP x H	<.01	0.971	<.01	1.00	0.321	0.02
Age x C	1.39	0.242	0.02	3.74	0.058	0.06
Age x C x AP	0.13	0.719	<.01	0.01	0.933	<.01
Age x C x H	0.01	0.903	<.01	0.26	0.610	<.01
Age x C x AP x H	1.32	0.255	0.02	0.10	0.753	<.01

Note: C= Condition; AP =Anterior-Posterior; H = Hemisphere; Age = Age Group. Significant results are in bold. Only effects that involve the condition factor are reported. N=62

The cluster-based permutation analysis however showed a significant cluster in the Younger Adults data (*p*=.015), with more negative ERPs for Prior FAs than Prior Slow CRs between approximately 690-830ms (as seen in Fig 5.15). This effect had a similar (although a bit more posterior) topography, latency and duration to the difference between Prior FAs and Prior Fast CRs. No significant clusters were observed in the Older Adults data, and this difference between Prior FA vs. Prior Slow CR was also significantly different between age groups (*p*<.001) around 710-830ms (Fig 5.15).

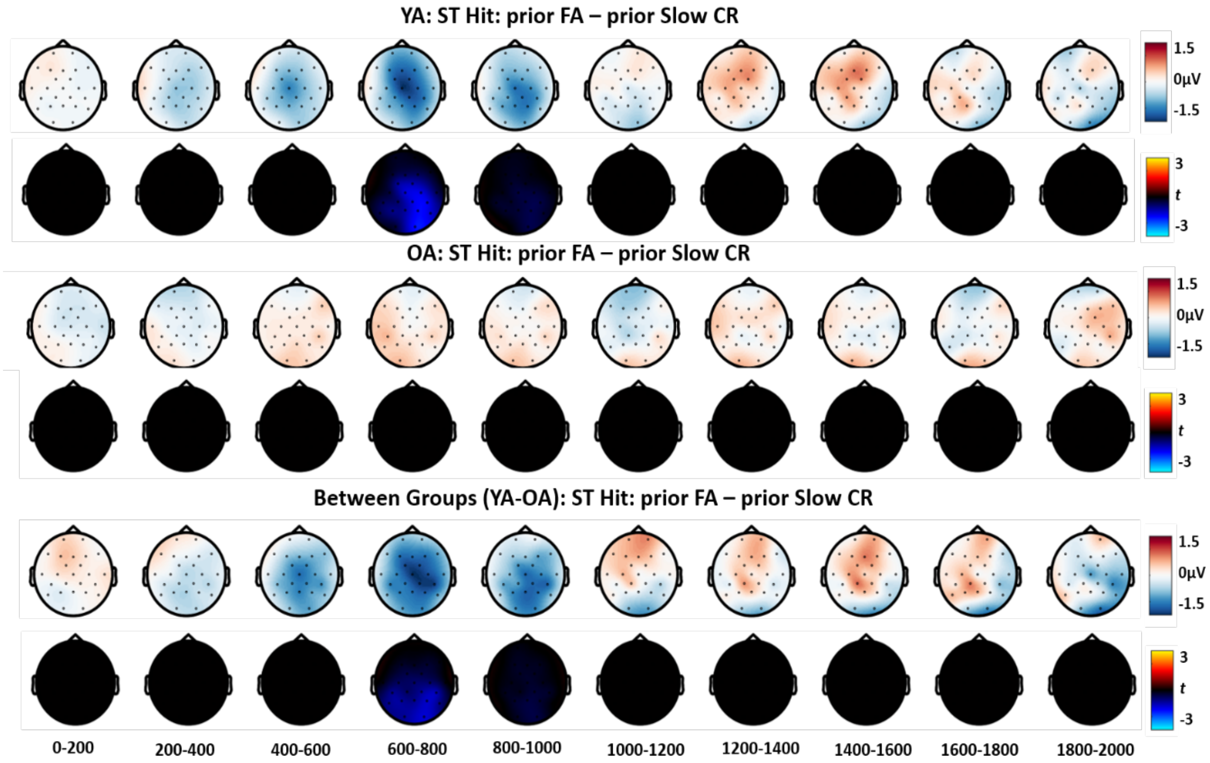


Figure 5.15. Whole-head analysis of ST Hits: prior FA v prior Slow CR ERPs. Topographical maps of amplitude differences (top rows, blue/white/red colour map) and t-values for the differences (bottom rows, cold/black/hot colour map) between ST Hits: prior FA v prior Slow CR split by Age Group and the final two rows showing the between groups difference. Only significant clusters (p < .025 at the cluster level) are shown in the t-statistic topographical maps.

ST Hit given that foils were a Prior Fast CR v Prior Slow CR

In the final ERP analysis of the surprise test data, I directly compared ERPs for Foil items that were correctly recognised on the surprise test (ST Hits) but previously received either a Fast CR (Prior Fast CR) or a Slow CR (Prior Slow CR) response on the first test. The purpose of this analysis was to investigate if the hypothesised novelty detection associated with Fast CRs on Test 1 resulted in different retrieval processing on Test 2 (such as enhanced recollection of the task context). Grand-average ERPs from the four electrode sites for each condition are shown in Figure 5.16.

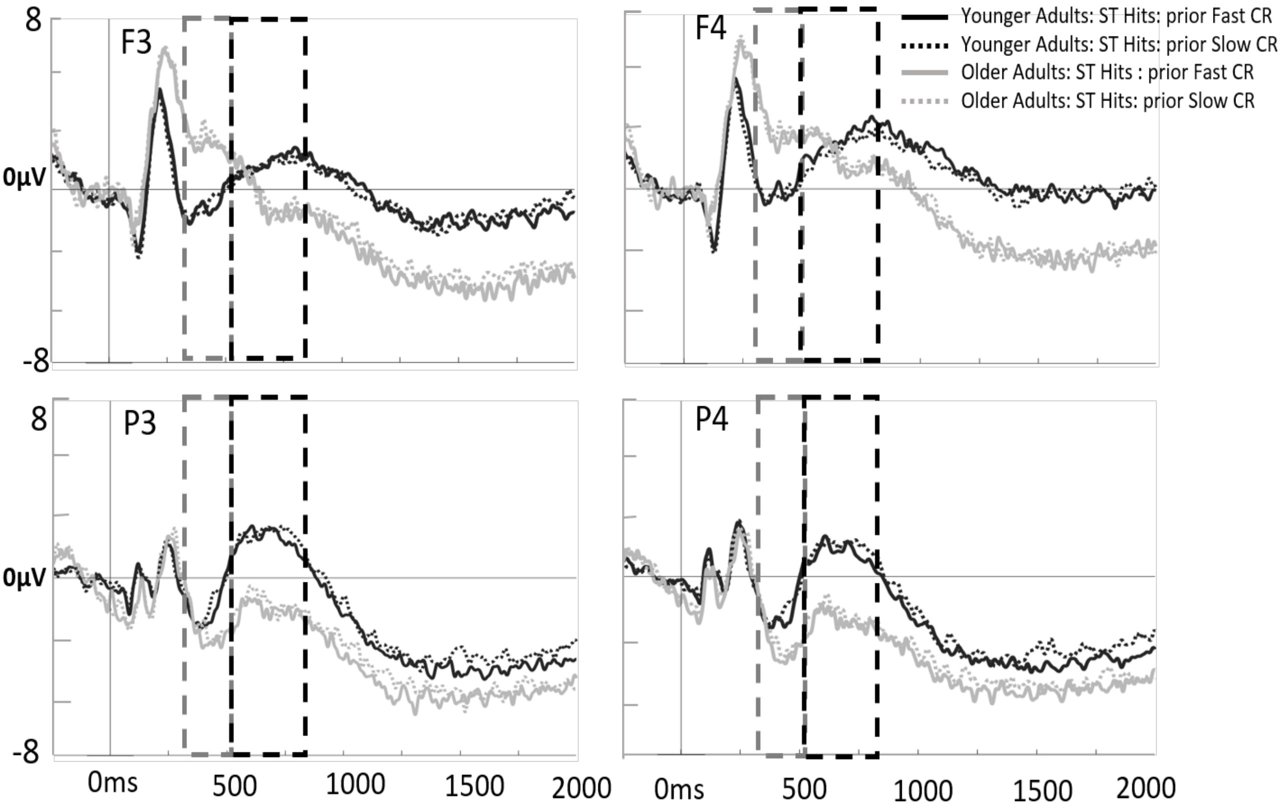


Figure 5.16. Grand-average ERPs from ST Hits; prior Fast CR and ST Hits; prior Slow CR from the Surprise Test for Younger and Older Adults at F3, F4, P3, P4. Boxes highlight the early 300-500ms and the late 500-800ms targeted time windows.

Results from the omnibus ANOVAs (factors; Condition: Prior Fast CR v Prior Slow CR; Anterior/Posterior; Hemisphere and Age Group) are shown in Table 5.26. Within the early time-window there were no significant interactions or main effects of interest. Within the later time-window there were also no significant effects, but there was a trend towards a 2-way interaction (C x AP).

Table 5.26. *Results from a mixed ANOVA for the omnibus test comparing Age effects on ERPs associated with ST Hits: prior Fast CR vs ST Hits: prior Slow CR during the Surprise Test in the two a-priori targeted time windows.*

Omnibus	Time Window					
	300-500ms			500-800ms		
	F	*p*	n_p^2	*F*	*p*	n_p^2
C	1.02	0.316	0.02	0.30	0.585	0.01
C x AP	0.82	0.369	0.01	3.19	0.079	0.05
C x H	0.56	0.458	0.01	0.25	0.622	<.01
C x AP x H	1.45	0.233	0.02	2.19	0.144	0.04
Age x C	0.03	0.874	<.01	0.77	0.383	0.01
Age x C x AP	0.07	0.794	0.00	0.85	0.359	0.01
Age x C x H	0.31	0.580	0.01	0.12	0.727	<.01
Age x C x AP x H	0.03	0.876	<.01	1.96	0.166	0.03

Note: C= Condition; AP =Anterior-Posterior; H = Hemisphere; Age = Age Group. Significant results are in bold. Only effects that involve the condition factor are reported. N=62

Likewise, there were no significant effects found in the cluster-based permutation test (Fig 5.17), either within Age Groups or between groups. The ERP results therefore suggest similar retrieval processes were engaged when foils were recognised on the surprise test if those foils had been correctly rejected as new in test 1, regardless of Test 1 response times.

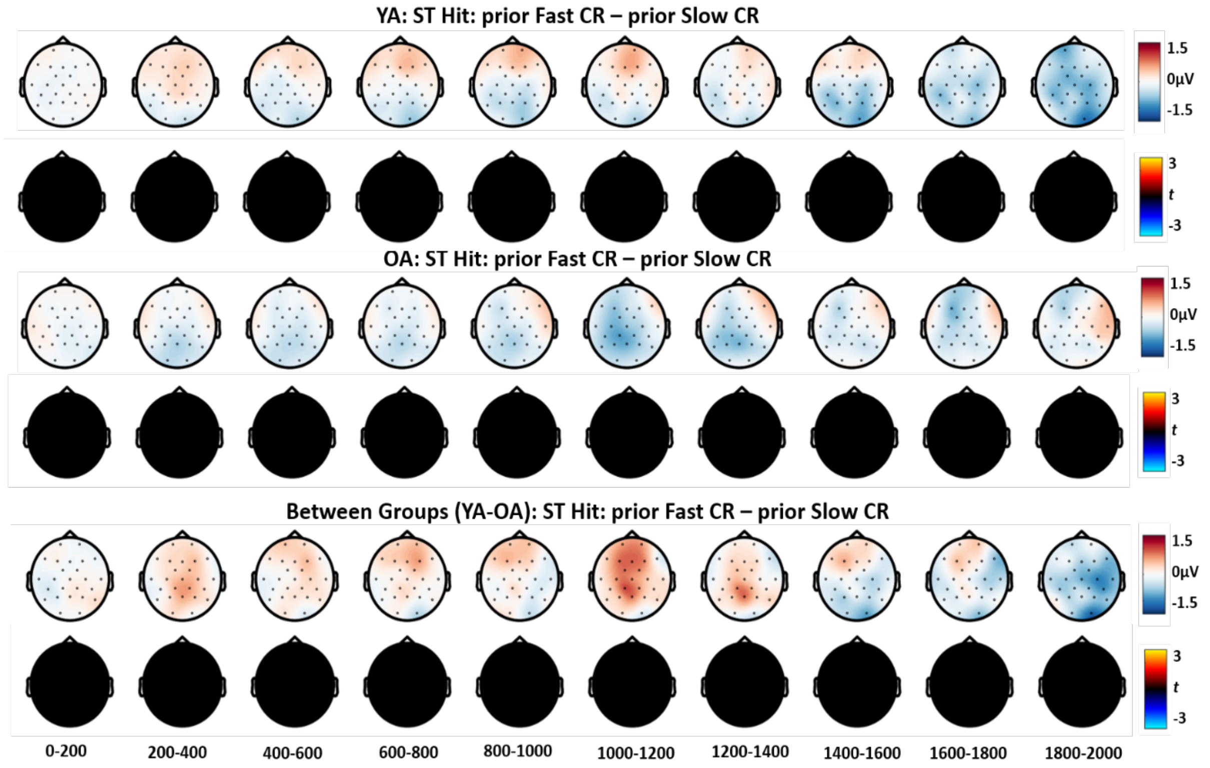


Figure 5.17. Whole-head analysis of ST Hits: prior Fast CR v prior Slow CR ERPs. Topographical maps of amplitude differences (top rows, blue/white/red colour map) and t-values for the differences (bottom rows, cold/black/hot colour map) between ST Hits: prior Fast CR v prior Slow CR split by Age Group and the final two rows showing the between groups difference. Only significant clusters ($p < .025$ at the cluster level) are shown in the t-statistic topographical maps.

Surprise Test Summary of ERP effects

Recognition related activity after incidental encoding during the initial recognition test showed predicted effects in the FN400 and parietal old/new effect time-windows, primarily in the younger group. This was investigated by comparing ERPs for all successfully recognised foils vs. new correct rejections, which revealed strong positive old>new differences in the FN400 and parietal effect time-windows in the younger group, but much-diminished amplitude differences in the Older Adults. The cluster analysis did reveal some significant late effects in the Older Adults, with increased positivity for CRs over hits. A similar but more widespread and sustained effect was also seen in the Younger Adults.

There were no early FN400 time-window differences relating to familiarity as a result of prior responses on Test 1, in contrast with my pre-registered prediction that foils that received false alarm responses on Test 1 would show a "carry-over" of familiarity-related ERP activity on the surprise test. However, unexpectedly, there was increased positivity in a later time-window (peaking ~600-800ms) for prior CRs (Fast and Slow) compared to prior FAs within the Younger Adults, with no significant differences in the Older Adults. No differences were seen between prior Fast and prior Slow CRs in either group.

5.4.4 Post Experimental Questionnaire Results

A post experimental questionnaire asked questions about strategy use, effort and motivation as well as if certain words were easier to recognition or remember. This was collected to provide some explorative insights into potential group differences.

Table 5.27. *Summary statistics for the post-Experimental Questionnaire.*

Experimental Phase	Questionnaire factors	Younger Adults	YA: N	Older Adults	OA: N
Study	Strategy used Y/N	96% (.19)	27	70% (.47)	30
	Effort during learning. 1-5	3.85 (0.77)	27	3.57 (.94)	30
	Motivation to win chocolate. 1-5	3.37 (1.28)	27	2.80 (1.45)	30
Recognition Test 1	Strategy used Y/N	48% (.51)	27	50% (.51)	30
	Effort during test. 1-5	3.96 (.66)	27	4.10 (1.00)	30
	Difficulty to recognise words. 1-5	3.11 (.75)	27	3.30 (.95)	30
	Certain words easier recognise. Y/N	100% (<.01)	26	93% (.25)	30
Surprise Test	Effort during test. 1-5	4.04 (.72)	26	4.30 (.79)	30
	Certain words easier recognise. Y/N	88% (.33)	25	79% (.41)	29
	Certain words easier to Remember. Y/N	63% (.49)	27	67% (.48)	30

Note: N differs due to some participants not responding to all questions. Y/N was a binary yes/no response and scores were coded 1 for yes and 0 for no. 1-5 was a scale where 5 represents highly effort / motivation or difficulty.

Independent t-tests indicated that there were no significant differences between age groups for effort, motivation or difficulty across all phases. The only significant difference was in strategy use during the encoding phase between the two age groups ($X^2(1)$ =6.79, phi=.35, *p*=.009) with younger adults reporting using strategies more often. No other significant differences were seen between strategy use during the test phase. Similarly, both groups equally reported that certain words in the recognition and surprise test were easier to recognise and remember compared to others. Participants reported that some words ‘stood’ out more and were inherently easier to identify.

5.5 Discussion

Experiment 4 was conducted to examine the electrical brain activity associated with recognition processing and decisions for novel items on an initial test and how those processes may influence subsequent recognition. I had directional hypotheses that more positive early FN400 time-widow effects would be associated with false familiarity (as indexed by false alarms to new items), whereas P3 time-windowed effects would be associated with heightened novelty detection (as indexed by rapid correct rejection of new items). ERPs recorded during the initial recognition test provided support for these two processes. I also had more explorative aims to examine these processes in older adults, to investigate age differences in retrieval processes and encoding of novel information during recognition. Some neural differences were seen across age groups, primarily with regards to successful recognition of old items when compared to new items, but there was also some evidence suggestive of differences in novelty processing in older adults. Additionally, as this paradigm utilised both intentional and incidental encoding, I expanded these project aims to explore how healthy ageing affected recognition processes after an intentional compared to an incidental encoding task. My ERP findings provide novel evidence that their neural processes associated with recognition are particularly reduced after incidental encoding in older adults.

As this project examined two main lines of enquiry, I will start this section by discussing the effects of intentional compared to incidental encoding on recognition processes across the age groups as these differences encompass more generalised memory processes. I will then discuss the effects seen in regard to specific differences between recognition processes for novel foils, and how this relates to incidental encoding. Finally, I will consider wider issues that may contribute to the similarities and differences seen across groups.

.

5.5.1 Age differences in recognition after intentional vs incidental encoding

Commonly there has been a focus in the memory literature on examining recognition performance and neural activation after intentional encoding, however incidental encoding is very important in everyday settings, and may be more widely used in older age as people are less likely to partake in formal education in later life (Wagnon et al., 2019). This project provided a good opportunity to explore ERP and behavioural recognition after incidental vs. intentional encoding across age groups.

Behaviourally, both age groups showed very similar recognition performance after the intentional encoding task on the first recognition test. Importantly, discrimination and response biases were very similar, within strikingly comparable patterns of RTs and Confidence ratings between conditions. This is unlike some recognition studies which suggest decreased discrimination ability in older adults (Fraundorf et al., 2019), with a more liberal response bias generally (Fraundorf et al., 2019; Suengas, Gallego-Largo, & Simón, 2010). The surprise test measured subsequent recognition of foils after incidental encoding that occurred as a by-product of the initial recognition judgment. On this test, younger adults showed better discrimination ability compared to older adults. Overall, both groups were more likely to guess 'Old' when uncertain, with the younger adults being the most liberal in terms of their bias. The difference in bias is more in line with prior findings that older adults have an increased tendency to guess 'Old' compared to younger adults (Fraundorf et al., 2019).

Given that behaviourally there were similarities across age groups in discrimination and bias after intentional encoding, it is reasonable to expect similarities in neural processes. The ERPs showed several similarities across the younger and older groups after intentional encoding. Younger adults showed a widespread early FN400 time-windowed old/new effect and later parietal old/new time-windowed effect when comparing hits vs. correct rejections on the first test. In the older adults, the early familiarity for old items was more focal to the right frontal scalp. This is interesting as prior findings have shown mixed effects, some consistent with these findings of intact early effects (Friedman, 2013; Nessler et al., 2007, 2008), to diminished (Duarte et al., 2006; Guillaume et al., 2009; Wolk et al., 2009) or even absent familiarity effects even though behaviourally familiarity processes are observed (Wang et al., 2012). It has been much debated what process within the FN400 or early mid-frontal time-windows indicates, with some arguing that this effect is more linked to semantic priming of items than familiarity (Paller et al., 2007; Voss & Federmeier, 2011). Both familiarity and semantic priming accounts could explain findings in this study, since the stimuli used were semantically meaningful words. The finding that both groups did show some evidence of these early ERP effects suggest that this process was engaged by both younger and older adults.

Similarly, whereas the later parietal old/new time-windowed effect was very large and had a typical left parietal peak in younger adults, this effect was also less widespread and was maximal over the frontal right region only in older adults (Finnigan et al., 2002; Li, Morcom, & Rugg, 2004; Rugg & Curran, 2007). Similar findings have been seen in prior studies, with suggestions being made that an additional central negativity ERP component might have diminished (through component overlap) the positive effects usually seen over the posterior and left regions in older adults, causing this apparent right lateralised effect. Tentative suggestions about causes of this central negativity have been linked how older adults retrieve information, with suggestions have been made that older adults employ a more visual or literal representation of a memory, whereas younger adults use a more abstract representation (Li et al., 2004; Rugg & Wilding, 2000). My questionnaire data indicated that younger adults were more likely to report using a specific strategy on the intentional encoding task than the older adults. If encoding strategies are reinstated at retrieval as suggested by the source-constrained retrieval literature (Dudukovic et al., 2009; Gray & Gallo, 2015; Jacoby, Shimizu, Daniels, et al., 2005; Shimizu & Jacoby, 2005), this could have created a difference in old/new retrieval effects due to the two groups spontaneously using distinctively different retrieval processes. However, my study was not designed to test for group differences in strategy type, and I only collected some limited qualitative self-reports about strategies used. Generally, of those that reported using a strategy and provided written information about the strategy, younger adults used more elaborative strategies, whereas there was a mixture of elaborate to less elaborate, shallow strategies in the older group. Therefore, I am unable to draw conclusions based on these reports, as further investigation is needed.

The results of whole brain cluster analyses generally supported the findings from the targeted analysis, but also revealed a much later difference between age groups, with opposing ERP effects from around 1650-1850ms, consistent with the timings and topography of the late frontal old/new effect (Hayama et al., 2008). Whereas the younger adults had more positive going ERP for hits than CRs in this time-window, the older adults showed opposing directional effects, with more positive going and right lateralised ERPs for CRs. The right frontal old/new effect is related to post-retrieval processes such as monitoring and evaluation (Hayama et al., 2008). Similar reversal of ERP activity between younger and older adults has been seen in other neuroimaging studies (Morcom, Li, & Rugg, 2007; Wolk et al., 2009) and suggestions have been made that older adults have deficits in pre-retrieval processes, and therefore rely more on post-retrieval monitoring and evaluation and using additional cortical areas to perform this controlled task. My findings thus support processing differences during recognition after intentional encoding, which could be due to differences in controlled processing and initial strategy use, or due to the engagement of compensatory mechanisms to increase retrieval accuracy through more effortful post-retrieval monitoring.

When comparing recognition processing differences across age groups after an incidental encoding task, there were more striking differences across the groups. This issue was investigated by comparing ERPs for prior foils that were successfully recognised on the surprise test with completely novel, correctly rejected items on the surprise test. On this test, younger adults showed a strong widespread and sustained old/new effect spanning the FN400 and left parietal time-windows, whereas older adults showed much-diminished differences between these conditions in the same time periods. This ERP result parallels with the behavioural findings that older adults had poorer recognition discrimination ability on the surprise test, suggesting impaired incidental encoding of foils. The lack of recollective neural activity for foils suggests that older adults should show large deficits in source judgments. However, this was not the case, since older adults performed well on the source judgment task for the foil items that they did recognise, and actually were more accurate than the younger adults, as younger adults misattributed prior FAs to the study phase more often. Thus, it appeared that recollection was not impaired for successfully encoded items in older adults, and since the ERP "hit" condition only included items that were successfully recognised, this result converges with prior studies showing absent recollection-related ERP activity but intact behavioural markers of recollection (Horne et al., 2020; Nessler et al., 2008). As discussed, the lack of activity could be linked to the centrally-maximal negativity seen in older adults at around the same time as the positive going partial old/new effect (Horne et al., 2020).

It is still not clear how age affects incidental encoding and the associated processes (Wagnon et al., 2019). There is evidence that incidental encoding is less affected by ageing than intentional memory uses, due to less reliance on controlled processes that are thought to be impaired by ageing (Old & Naveh-Benjamin, 2008; Téllez-Alanís & Cansino, 2004). Larger age differences in memory are seen when the tasks require cognitive control and elaborate processing (Wagnon et al., 2019). Therefore, in this study, age differences in recognition memory for incidentally encoded foils could have occurred if such encoding depended on the degree to which participants used strategic control during recognition attempts. That is, encoding of foils during recognition might be more dependent on control processes than other types of incidental encoding tasks, that typically direct participants to a certain judgment type, such as incidental encoding as a by-product of a deep judgment task. This result thus provides interesting novel evidence about age differences in recognition-related ERPs after incidental encoding, when encoding is occurring during recognition attempts. This finding is important as older adults will often rely on incidentally encoded information, which we know little about from research (Wagnon et al., 2019), but may explain the day-to-day complaints of poor memory from older adults.

The whole scalp cluster analysis revealed widespread effects in the later part of the epoch with increased positivity for correctly rejected new items compared to hits (prior foils) on the surprise test. Similar effects were present in both groups, but the difference was somewhat stronger and more sustained in the younger adults. This finding is interesting as the younger adults here showed consistent directional effects in line with the older adults, whereas after intentional encoding the younger adults showed an opposite directional late ERP effect. These ERP patterns may relate to differences between the first recognition test and the final surprise test in terms of encoding and/or retrieval demands, since the surprise test also included a source judgment task after old/new recognition, which may have increased task demands. This is interesting as it appears that post-retrieval processes are very much intact in older adults and the differences seen may be reflecting the level of usage and efficacy (Horne et al., 2020; Wang et al., 2016; Wolk et al., 2009).

Therefore, there appear to be neural similarities between recognition related neural activity after intentional encoding in younger and older adults, but with some smaller changes. After incidental encoding, larger age differences are seen with diminished/absent recognition-related ERP activity for recognition of foils in older adults. In the next section, I discuss the effects of item-specific retrieval and decision processes on incidental encoding of foils, which might shed more light on the processes that could have caused these age differences in ERP activity on the surprise test.

5.5.2 Recognition processes that relate to encoding during retrieval attempts: familiarity and novelty

Of particular interest in this study were the neurocognitive processes associated with increased recognition of certain foils over other. Behaviourally, I saw increased final recognition for foils both if they were associated with prior false alarms and if they were rapidly identified as new, suggesting potentially two distinct processes occurring during recognition judgments. This pattern was found when including all participants and trials, replicating the findings from Experiment 3. It's important to note however that when examining only the ERP sample and including only ERP trials after separating the foils by accuracy and response time on test 1, some changes in performance were observed (likely introduced because the ERP sample inclusion criteria was performance-dependent, thus biasing behaviour in this sub-group). The younger adults no longer showed subsequent boosts in recognition discrimination for the foils that received prior false alarms over correct rejections (even though behaviourally this was seen in the pilot and full sample). There was an increase in final test discrimination however for the foils that received prior fast over slow correct rejections. In the ERP sample, the older adults showed effects more in line with predictions, with increase subsequent recognition for foils that received prior false alarms over correct rejections, with contrary to predictions, slow correct rejections being recognised more the fast correct rejections. Yet surprisingly no subsequent boost in recognition of foils that received prior fast over slow correct rejections was seen in the older group.

5.5.2.1 False Familiarity

Enhanced recognition of foils that received false alarm responses on test 1 could have been driven by a persistent false memory for those items, or enhanced incidental encoding after making an error. Even though item discrimination measures showed little differences between different foil types in the younger group, source judgments provided insights about the underlying processes. Both groups made more incorrect source judgements on the final test for prior FAs than prior CRs, misremembering seeing those items in the study phase rather than encountering them for the first time in the test phase. There was no difference in source memory for the test phase across foil conditions in either group, suggesting that both age groups were susceptible to persistent false memories that carried over from the first to second test, rather than enhanced encoding of foils as a result of error-related processing (e.g. due to enhanced attention to errors or similar). Surprisingly, younger adults actually showed stronger false source memory for seeing prior FA foils in the study phase. This group difference was unexpected since older adults are argued to have impaired recollection and cognitive control processes, known mechanisms involved in making correct source judgments (Fraundorf et al., 2019; May, Rahhal, Berry, & Leighton, 2005).

Overall, the source judgments provided support that the FA and CR foil conditions were processed in different ways for both groups, regardless of the limited recognition discrimination differences between these conditions in younger adults (which may have been caused by ceiling effects on the old/new recognition task). The behavioural responses overall indicate very similar performance across the two age groups (and especially in the first test), which is useful for interpreting potential neural differences that will be less confounded by memory differences that are related to performance rather than age.

In line with predictions and the behavioural findings, on the first test, foils associated with FA responses (FAs in short) elicited reduced ERP negativities in the early time-window compared to foils associated with fast and slow correct rejections (Fast CRs and Slow CRs in short), although this effect was strongest for the latter comparison. This early decreased negativity is compatible with an early 'FN400 like' time-windowed effect (Friedman & Johnson, 2000; Mecklinger & Jäger, 2012; Rugg & Curran, 2007), which is related to familiarity processing. This pattern is consistent with predictions and the behavioural responses as a high proportion of FAs elicited a false source memory of studying those items. Both the ERP and the behavioural data therefore converge to support the suggestion that FAs hold more false familiarity and related false source memories, compared to the new items. This conclusion is consistent with prior literature showing early 'FN400 like' time-windowed old/new effects also for FAs and not only Hits (Wolk et al., 2006), suggesting this effect is related to the subjective experience of recognition rather objective accuracy of recognition decisions. Interestingly, this effect was not significantly affected by age, with both age groups showing similar effects. It is worth noting that the cluster analysis which corrects for multiple comparisons is more conservative than the targeted time-window analysis, and therefore did not detect this effect as significant.

This study provides a novel overlap between the ERP ageing literature on familiarity processes as well as the false memory literature. The intact early 'FN400 like' time-windowed effect for FAs in older adults adds to the ERP literature relating to healthy older adults, as currently there is limited research investigating these early effects with mixed findings so far (Duarte et al., 2006; Friedman, 2013; Guillaume et al., 2009; Nessler et al., 2007, 2008; Wolk et al., 2009) even though familiarity is often claimed to be unchanged by ageing. The findings here are consistent with studies showing intact familiarity-related ERP effects in older age (Friedman, 2013; Nessler et al., 2007, 2008), and suggests that similar processes are elicited by falsely recognised stimuli in younger and older people. However, there are still debates over what the early ‘FN400 like’ time-windowed effect is actually measuring (Paller et al., 2007) which I will consider in more detail in the general discussion in the next chapter, together with suggestions to further examine familiarity effects.

Cluster analysis also revealed a late positive going ERP effect for FAs compared to Fast and Slow CRs within the older adults, whereas younger adults shared this directional effect but it was much weaker. As discussed in relation to recognition after intentional encoding, late ERP slow-drifts may index the engagement of post-retrieval processes, suggesting that there were differences in the evaluation and monitoring of the FAs compared to CRs. Potentially since the FAs are identified incorrectly, they also underwent more deliberation and monitoring, with both groups showing this directional effect but the older adults potentially relying on these post-retrieval processes more heavily as compensatory mechanism (Wolk et al., 2009). These ERP effects could also be linked to the older adult’s self-efficacy beliefs, as overall the older adults responded with lower confidence in their recognition responses, especially for FAs. Lower self-efficacy has been reported as a potential cause of lower memory performance due to beliefs that older adults have worse memory (Berry et al., 2010; Fraundorf et al., 2019; West et al., 2009).

5.5.2.2 Novelty processing

Somewhat later than early ‘FN400 like’ time-windowed effect were strong ERP differences in the ‘P3 like’ time-windowed responses that may relate to novelty detection. Such effects were expected to be driving the separation between rapidly detected new items (Fast CRs) compared to items more slowly detected as new (Slow CRs) as well as FAs. Previous research has shown that items that are more novel in a context are incidentally encoded at a high rate as these capture attention (Barry et al., 2020). Both age groups showed an increased P3-like positivity for Fast CRs compared to FAs. The younger adults had a very widespread topography of this effect, whereas the older adults showed a broad distribution except the front-right scalp locations. This evidence for an increase in novelty processing was also associated with increased confidence in response accuracy for Fast CRs. Confidence responses have been linked to the strength of other processes such as the parietal old/new effect (Finnigan et al., 2002) and therefore a similar graded response could be occurring here but in relation to novelty strength.

Unexpectedly, in the older adults, Slow CRs also showed increased novelty-related effects in the P3 time-windows in the front-left scalp region. Slow CRs were not expected to induce a heightened novelty signal, as seen in the younger adults where FAs and Slow CRs had similar neural activity in this later time-window, as I suggest these are discriminated more based on the familiarity of items, rather than the novelty of the CRs. The pattern of results between Fast and Slow CRs supports this view only for the Younger Adults, who had an ERP difference in the P3 time-windows, with ERPs for Fast CRs being more positive than Slow CRs. Older adults however, had no ERP difference between these conditions, indicating that potentially all CRs were processed in a similar way. This lack of neural difference is consistent with the lack of behavioural difference in subsequent foil recognition accuracy between these conditions in the older group, with older adults in fact having higher subsequent recognition of Slow CRs. This is in direct opposition to the prediction that Fast CRs would be more novel and thus be incidentally encoded more often. For older adults, there was also no difference in source response confidence for these conditions, whereas younger adults were more confidence about the source judgment for Fast CRs.

A potential explanation for these results is related to the increased processing time of the Slow CRs in the older adults, which may have increased the engagement with these items and the likelihood of subsequent recognition. This proposal could also be coupled with the idea that older adults may be using novelty processing more generally than younger adults. Novelty processing has been argued to remain intact over age (Berti et al., 2017; Daffner et al., 2005; Friedman et al., 1998; Richardson et al., 2011). Therefore, it could be tentatively suggested from these findings that older adults strongly rely on novelty processing in their recognition decisions, helping them maintain good recognition performance for new items, but that doing so is not solely causing the increase in incidental encoding. The engagement with items also seems to be driving and interacting with these novelty signals, with more processing time linking to more deliberation and more incidental encoding. Since limited research has been conducted investigating the novelty P3 in healthy older adults, these findings provide important new insights and support the view that older adults have intact novelty processing (Berti et al., 2017). Similarly to how the parietal old/new effect amplitude is linked to recollection strength (Rugg & Allan, 1995; Vilberg et al., 2006; Wilding, 2000), it would be interesting to examine if the effects and amplitude in the P3 time-window is also moderated by novelty amount. In younger adults there is a clear habituation in the P3 time-window responses with repeated exposure (Barry et al., 2020), with items that are less novel showing a reduced effect. Therefore, it does seem that younger adults do vary their novelty signal based on the stimuli novelty strength, seen here and in prior literature (Barry et al., 2020; Debener et al., 2005; Dinteren et al., 2014; Friedman et al., 2001). Yet, as seen here older adults show a constant similar novelty processing for all correct rejections, supported somewhat by prior literature showing a the reduce habituation of novelty processing after repetition (Friedman et al., 1998; Richardson et al., 2011). Testing this further would provide clearer insights into how novelty is processed in older adults and how novelty contributed to incidental encoding.

5.5.2.3 Subsequent Successful Recognition

On the surprise test, I compared ERPs during correct recognition of foils after different types of initial item processing on Test 1, specifically for foils that were previously falsely classified as old in Test 1 (Prior FAs), and foils previously quickly (prior Fast CRs) or slowly (prior Slow CRs) correctly rejected as new on Test 1.

I predicted that on the surprise test, false familiarity for FAs on Test 1 will 'carry over' to the final test, so that ERPs during the surprise test will show enhanced early (FN400 time-window) positivity for previous FAs compared to previous CR conditions. However here, no early familiarity related differences were seen as a result of prior responses on Test 1. I also predicted that if false familiarity on Test 1 enhances the encoding of context, then previous FAs will also be associated with enhanced ERP markers of recollection (parietal old/new effect) compared to previous CRs. However, unexpectedly, there was increased positivity in a later time-window (peaking ~600-800ms) for prior CRs (Fast and Slow) compared to prior FAs within the Younger Adults, with no significant differences in the Older Adults as opposed to expected increases in recollective activity related to FAs. Consistent with predictions however, source responses were different for prior FAs and prior CRs, with participants giving more accurate responses for previous CRs than previous FAs, with FAs being falsely judged as recognised from both study and test phases (even though they were only shown in the test) whereas previous CRs should be primarily associated with source memory for the test phase.

Moreover, I predicted that if novelty detection on Test 1 enhances the encoding of context, then previous fast CRs will be associated with enhanced ERP markers of recollection (parietal old/new effect) compared to previous slow CRs. These results showed the only difference in ERPs associated with foil recognition activation were within the younger adults, who showed an enhanced ERP positivity for prior Fast CRs compared to prior FAs. A weaker but similar effect was seen with increased ERP positivity for prior Slow CRs than FAs, again only within the younger adults. The ERP positivity associated with CRs emerged in a similar time-window as typical recollection-related ERP effects, but had a more central peak. Despite this difference in topography, I am tentatively interpreting this effect as consistent with increased subsequent recollection-processing for foil items that were detected as novel during incidental encoding. This view is consistent with findings that the strength of the parietal old/new effects is graded with recollection strength, and can therefore vary across old items that are successfully recognised (Vilberg et al., 2006). This claim is also supported by the finding that young adults showed increased confidence in correct 'Test' source responses for Fast CRs, followed by Slow CRs even though there was not a boost in source accuracy, as previously predicted. This finding also converges with my previous interpretation that the increases in final recognition for prior FAs is not related to the increased encoding of these items in the test context, as previously predicted, but rather due to a sustained false memory for those items, either due to semantic relatedness of the items within the word lists, or due to extra-experimental familiarity of items on an individual level.

It is not surprising that this activity pattern is only seen for the younger adults, as the overall retrieval-related ERP activity was diminished in the older adults during recognition after the incidental encoding task, as compared to the more intact activity after the intentional encoding task. This finding however does raise issues with both what these components are measuring, as well as about how ERPs are calculated, given that subsequent hit rates for older adults was still very high and source accuracy was in line with that of the younger adults. These theoretical and methodological issues will be delved into more in the general discussion.

5.5.2.4 Wider considerations

There are wider considerations to deliberate on given that the older adult participants were mostly (with some exceptions) highly educated and socially, physically active. These are known lifestyle factors that protect against memory deterioration (Chan et al., 2018; Shafto et al., 2019). My results therefore might not be representative of the whole population of healthy older adults, but instead of a subgroup that have self-selected to take part in psychology studies. It could be argued however that this type of sub-group is more similar to the student populations that are also highly educated and generally socially and physically active, and thus have more intact cognitive functions. Emerging research has shown that highly performing older adults show more intact cognitive processes and more similar neural activity to young adults, as compared to low performing older adults (de Chastelaine et al., 2016; Duarte et al., 2006; Wang et al., 2016). This study therefore could have primarily included 'high performers', yet because of the relatively small sample size, it is not possible to investigate performance effects on ERPs within this study. Future research should examine if even within a high performing population there are correlations between lifestyle factors and memory function.

In this study there were no instructions given to direct use of strategies or the type of strategy to use, but participants were explicitly told that they would be tested on the words they encountered in the first study phase. I did provide a constant small incentive to try to maintain task engagement, and there seemed to be similar reports of effort and motivation during the tasks across groups. However, age and individual variations in strategy use could have occurred, which may be a particular concern in the older adult group as older adults may be less likely to self-initiate strategy use (Morcom, 2016; Saczynski et al., 2007), and if they do, they may be more likely to use non-elaborate strategies which are less effective (Kirchhoff et al., 2012). From the self-reports, it is clear that both groups self-initiated strategy uses during the initial study phase, with nearly all younger adults (96%) using a strategy, but with only

70% of older adults using a strategy. Even though fewer older than younger adults self-initiated a strategy, a very high proportion did, which could have helped maintain the high recognition rates on the first test in the older group. Research has found that even though older adults use strategies less often, they are still able to effectively use strategies if they chose to do so (Castel et al., 2011), especially when they have generated an elaborate strategy (Flegal & Lustig, 2016). One can speculate that highly educated and socially active older adults might be particularly prone to use memory strategies to aid their busy lifestyles. This is conjecture, but it would be an interesting line to investigate further, in particular the link between strategy use in older adults and intentional encoding as well as during retrieval attempts that affect incidental encoding.

An important but difficult factor to control is potential 'jitter' (temporal variability) that may affect older adults' ERPs. ERPs are averaged single time-locked potentials that rely on consistent timing of activity over trials. The process of averaging across trials will reduce the size of ERP effects that are not time-locked or that are slightly 'jittered' in time. Older adults' neural activity may be more jittered in time (Cabeza, Nyberg, & Park, 2009; Picton et al., 2000) which created issues when using average ERPs to analysis group differences. This makes it difficult to compare and contrast age group ERP activity due to the variability seen in older adults of ERP components that are traditionally very robust in younger adults. This is a wider issue relating to ERP use with older adults which I discuss briefly in Chapter 6.

More generally, considerations also need to be made when analysing ERPs as they require a certain amount of trials to produce a reliable average ERP (Friedman & Johnson, 2000; Picton et al., 2000; Rugg & Allan, 1995). Because I designed this experiment to analyse ERPs for error trials (false alarms), this therefore led to the exclusion of very high performers as the conditions of interest were dependent on the individual performance on test 1. There were similar numbers of top performers excluded from both the younger (*N*=7) and the older adult groups (*N*=9). This issue is therefore less of a concern for between-group findings, and is primarily a theoretical issue about what processes are likely to occur in participant who have low FA trial numbers compared to those who have high FA trial numbers. Linking these effects to predictions from signal-detection theory (SDT; Wixted, 2007) it is reasonable to expect participants that have high FA numbers to rate these with lower confidence, relying more on guesses due to a more lax response criterion due to more overlap between old items classified as new and new items classified as old. Compared to those that have less FA trials, with these trials being more likely linked to stronger false-memory related errors having a much more conservative response bias and stricter criterion with less overlap between the distribution of responses related to recognising old and new items. Therefore, following on from this, I would suggest that different processes would underlie those that make false alarm errors due to false memory related activity, and those that are due to more deliberation and evaluation of familiarity when guessing. I would expect to see larger effects associated with false familiarity for participants who make fewer FA responses based on this suggestion. Behaviourally I did see a reduction in subsequent differences when including all trials compared to ERP only trials, which starts to support this suggestion. However, it is very difficult to compare participants based on high and low numbers of FAs as those with few FAs could have very noisy or unrepresentative ERPs for FAs. Due to this issue, the decision was made to use the laxest trial threshold possible (N>14 trials) for inclusion in ERP sample.

5.6 Conclusions

The present findings provide evidence for multiple distinct neural processes occurring during recognition judgments, that relate to encoding of novel information in different ways. Interestingly, these processes occur even without explicit instructions or task manipulations, providing support for the item-specific spontaneous nature of these processes. It has been further shown here, in line with other memory-for-foils literature (Alban & Kelley, 2012; Bergström et al., 2015; Dudukovic et al., 2009; Jacoby, Shimizu, Daniels, et al., 2005) that recognition attempts also result in incidental encoding of new items. Here I show that incidental encoding of such items relates to the retrieval and decision processes involved during the initial recognition judgement. Not only do familiarity and recollective processes occur, but I also found evidence for a distinct novelty-related process, in line with literature about the P3 component as a marker of novelty detection (Barry et al., 2020; Friedman et al., 2001). The findings from this study supports my pre-registered hypotheses that there are at least two distinct processes occurring during recognition attempts of new foil items, familiarity and novelty detection.

Interestingly, older adults show similar familiarity ERP effects associated with false alarms, but differences to younger adults in novelty ERPs as a function of speed of decisions, suggesting that the older adults may use novelty during recognition in a different way than younger adults. Additional age differences were seen in the surprise test, with diminished ERP activity related to recollection in older age, even though source judgments were more similar across groups. Interestingly, whole-head cluster analysis shows unpredicted later increases in post-retrieval monitoring in older adults, which may be responsible for the behaviourally preserved recognition. Finally, within the younger group only, foils that were associated with novelty detection when initially encountered in test 1 elicited ERPs consistent with increases in subsequent recollection, showing the after effects of novelty detection on subsequent memory. This study therefore raises questions about the causes and implications of false familiarity. Other questions are raised about the effect of differences in novelty processing in older age, and how this impact incidental encoding.

5.7 Chapter Summary

Overall, this chapter brings together multiple findings from the thesis and directly tested questions raised in Chapter 4 by investigating what electrical brain activity occurs during a recognition attempt that leads to differences in incidental encoding. This study also extends on my specific hypotheses to explore memory-related ERP activity in older adults, which is associated with mixed, contradictory findings in the literature. Overall, my findings provide novel additions to the currently inconclusive literature about recognition processes in older adults. In the next chapter, I discuss the findings from the whole thesis and draw conclusions based on the whole body of work. I also make critical suggestions based on new knowledge as well as provide my reflection on the limitations in the current studies.

Chapter 6: General discussion, conclusions and suggestions

This thesis examined novel factors that could potentially moderate the rate of incidental encoding during recognition attempts. The factors examined are known to affect episodic memory more broadly. This research question was motivated by the Source-Constrained Retrieval (SCR) literature (Alban, 2013; Alban & Kelley, 2012; Bergström, Vogelsang, Benoit, & Simons, 2015; Danckert et al., 2011; Gray & Gallo, 2015; Halamish, Goldsmith, & Jacoby, 2012; Jacoby, Shimizu, Daniels, et al., 2005; Marsh et al., 2009; Stark & Okado, 2003), which has widely used the memory-for-foils paradigm. As previously shown, the rate of incidental encoding of foils in this paradigm is modified by the reinstatement of initial encoding strategy or task, with evidence from a broad range of factors; self-referential thinking (Bergström et al., 2015) survival ratings (Nairne et al., 2015), as well as the widely tested Levels-of-Processing (LOP) manipulations (Jacoby, Shimizu, Daniels, et al., 2005). Other studies have shown effects of full versus divided attention at recognition as modulating encoding of foils (Dudukovic et al., 2009). This literature has thus demonstrated that the particular processes and strategies that are engaged during retrieval attempts have a knock-on effect on the rate of incidental encoding of novel information, as well as showing that episodic memory tasks typically characterised as encoding vs. retrieval tasks are not 'process pure' (Richardson-Klavehn & Bjork, 1988).

An important factor that I wanted to examine was how incidental encoding during recognition attempts is affected by age, with only one previous study examining this issue (Jacoby, Shimizu, Velanova, et al., 2005). Episodic memory functions have often been observed to decline with healthy ageing (Craik, 1994), yet more recent evidence suggests such decline does not always occur, and a more holistic take is needed (Henson et al., 2016; Josefsson et al., 2012; Shafto et al., 2019). This thesis explored what potential factors could boost the rate of incidental encoding of foils, exploring the effects of external reward, item-specific retrieval and decision processes, as well as how these factors are affected by healthy ageing. I also examined the neural correlates of these processes by measuring electrical scalp activity with ERPs, therefore providing a novel take on the SCR literature which has primarily been examined with behavioural and fMRI methodology so far and with young populations (Bergström, Williams, Bhula, & Sharma, 2016; Jacoby, Shimizu, Daniels, et al., 2005; Nairne et al., 2015; Stark & Okado, 2003).

In this closing chapter of my thesis, I will consider more broadly the effects I found in relation to the prior literature on incidental encoding during recognition attempts. I will start by providing an overview of the key findings from my empirical chapters, then discuss the theoretical implications of the thesis. I will then move on to discuss the limitations and provide future directions based on the findings presented here.

6.1 Summary of empirical findings

In **Experiment 1 in Chapter 2** I aimed to replicate the one prior study to my knowledge that had examined SCR in older adults (Jacoby, Shimizu, Velanova, et al., 2005) where they found ageing-related deficits in spontaneous reinstatement of LOP encoding strategies. I expected to see similar ageing-related deficits in the SCR effect, but I also expanded on this research to explore the potential modulating effect of external rewards (in this case, monetary), where I expected to see that rewards enhanced incidental encoding of novel foils. Therefore, in Experiment 1 (*N*=82) the memory-for foils paradigm was modified and Younger and Older Adults completed a SCR task investigating both LOP and Reward effects on intentional and incidental encoding during retrieval. This extended the literature on LOP and SCR, and contributed a novel investigation of the effects of Reward. The key findings were, unlike previously suggested, that older adults could successfully spontaneously and strategically constrain their memory search as I showed similarly matched patterns of the LOP SCR effect in both groups. This result was surprising, but suggests that ageing-related decline in strategic retrieval processes is not universal and wider considerations are needed, such as lifestyle factors (Chan et al., 2018; Nyberg & Pudas, 2019). My findings conflicted with literature on reduced self-initiated strategy use in older adults (Morcom, 2016), but fits with emerging literature showing intact strategy implementation in older adults (Flegal & Lustig, 2016; Kirchhoff et al., 2012). When examining the novel factor of rewards, it was apparent that only the younger adults responded positively to this manipulation. In the young age group, recognition performance was enhanced after intentional encoding when rewarded, but no apparent difference due to reward was seen after incidental encoding. I concluded that potentially the reward manipulation was too subtle and not able to influence recognition processing enough to modulate incidental encoding rates.

Therefore, extending on the findings in Experiment 1, in **Experiment 2 in Chapter 3** I attempted to more directly and specifically manipulate recognition processing through altering the relative value of different items during the recognition test, with points representing a potential monetary reward. This experiment was very explorative given the novel paradigm which combined both the memory-for-foils structure (Jacoby, Shimizu, Daniels, et al., 2005) and the value-directed remembering point manipulation (Castel, 2007), I only recruited younger participants to take part (*N*=86), since only the younger adults in Experiment 1 had shown evidence of different recognition memory as a function of reward. The memory-for-foils procedure was modified to no longer use a blocked design, and use an intermixed conditional manipulation. Here, point values were introduced during the recognition phases, with some old and new words being of high value and others of low value. The key findings from this experiment were that participants altered their retrieval processes due to point values that were presented during recognition. Both response times and response bias were altered due to point value, suggesting a prioritisation of processing of the high value items. However, even with these differences in recognition processes there was no alteration in subsequent recognition of foils, suggesting no effect on incidental encoding during those recognition attempts. This is in striking contrast to the widely seen effects of point values on intentional encoding (Castel et al., 2007, 2011, 2013; Hennessee et al., 2017, 2019; Middlebrooks & Castel, 2018), suggesting that incidental encoding during recognition tests is less sensitive to such manipulations, at least when final foil memory is assessed with recognition rather than free recall as typically used in the value directed remembering literature.

The strongest effects of reward on recognition processes were seen in Experiment 1 and therefore in **Experiment 3, in Chapter 4** I aimed to re-examine the effects of rewards when presented during intentional encoding. I also started examining the effects of subjective recognition experiences and item-specific retrieval and decision processes that might affect incidental encoding of foils in the memory-for-foil paradigm. Subjective recognition judgments were assessed by continuous confidence ratings during the first recognition test, and feedback was provided to increase the saliency of reward, confidence, and decision accuracy factors, in order to build on the hypercorrection literature which has shown enhanced encoding of new information in response to feedback about high confidence errors (Butterfield & Metcalfe, 2006). The Remember-Know (Donaldson, 1996; Gardiner & Richardson-Klavehn, 2000) scale was used during the final foil recognition test, to provide a more informative measure of whether the factors under investigation resulted in recollective and/or familiarity-based subsequent recognition of foils. Again, as this was an explorative research question, I focussed this study on potential effects in younger adults (*N*=48). Similarly, to in Experiment 1 and 2, there were no effects of reward value on the rate of incidental foil encoding, yet interesting findings were revealed that were related to how recognition decision accuracy for foils during the first encounter effected subsequent recognition of those foils. Specifically, two different effects were found whereby foils that received false alarm recognition responses or very rapid correct rejection responses on Test 1 were more likely to be recognised as old on a final test. This data pattern therefore suggested that false familiarity and novelty detection processes might both be important modulators of foil encoding. However, as the task was not optimised to investigate these effects, I was not able to conclusively determine the underlying mechanisms that enhanced foil memory for these types of items. Furthermore, since these patterns of results were found after a large number of exploratory analyses, I could not be confident that the results were replicable.

Therefore, the next study was designed to replicate the results from Experiment 3, and to elucidate the underlying neurocognitive mechanisms that gave rise to enhanced foil recognition after false alarm errors and rapid novelty detection. I therefore developed a suitable version of the memory-for-foil paradigm to use with ERPs, which was piloted (*N*=17) to verify that the improved design was successful at replicating the behavioural pattern. In **Chapter 5, I present Experiment 4,** which directly address these issues. In this study, I recorded ERPs in younger and older adults during both the initial recognition test and the final surprise foil tests, and added source judgments to the final foil test. These judgements were added to more directly examine whether foils were recognised on the final test because they were falsely remembered from the study phase, or because they were accurately remembered from the test phase. Key results from the pilot replicated the findings from Experiment 3 that foils that received false alarm responses on Test 1 were more likely to be recognised as old on a final test, and the source judgments indicated that persistent false memories that carried over from the first to second test were likely driving this result.

As piloting had determined that my improved paradigm was successfully replicating and extending behavioural findings from Experiment 3, this version was then used with ERPs to examine what neural processes were underlying these behavioural memory effects. My pre-registered predictions were that there should be an increase in early familiarity ERP effects (FN400) for foils receiving false alarm errors, whereas I expected to see a later increase in ERP positivity (Novelty P3) for Fast compared to Slow CRs if these are associated with novelty detection. **In Experiment 4** (*N*=62), I not only examined my pre-registered predictions for young adults, but I also extended these aims to re-examine memory-for-foils and underlying neural processes in healthy older adults, linking to the effects in Experiment 1. I therefore investigated whether Older Adults would have similar memory performance to the Younger Adults (as seen in Experiment 1), and whether such similar behaviour would be reflected in similar or different (e.g. Friedman, 2013) neural activity.

Critically in Experiment 4, the ERP findings were mostly consistent with predictions, with increased familiarity-related 'FN400' time-windowed effects for foils associated with false alarm errors, and increased novelty-related 'P3' time-windowed effects for foils associated with fast correct rejections during recognition test 1. Behavioural results also overall converged with the claim that in younger adults, there are at least two distinct processes occurring when attempting to recognise novel items, and that recognition decisions are not just based on familiarity strength, but also based on novelty strength. Experimentally novel foils that elicit neural evidence of familiarity are also associated with a false memory experience that is reinstated across repeated recognition attempts. In contrasts, foils associated with heightened novelty are more strongly incidentally encoded, perhaps due to increased attention allocated to these items (Ranganath & Rainer, 2003).

Older adults similarly showed the FN400 time-windowed familiarity-related effect for foils associated with false alarms but did not show a difference in novelty related P3 time-windowed effects between fast and slow correct rejections, with a similarly strong novelty related P3 time-windowed effects for all such items. This ERP finding converged with an increase in subsequent recognition and source memory of both types of correct rejections in the older group, suggesting that older adults rely on both familiarity and novelty processing during recognition decisions, but in a different way to younger adults. Unlike in younger adults, processing time for foils seems to benefit older adults by increasing incidental encoding rates.

Additional findings in Experiment 4 also provided novel insights into recognition processing after intentional compared to incidental encoding in healthy older adults. I found that older adults had somewhat altered but still evident neural activity associated with recognition after intentional encoding, consistent with prior findings (Li et al., 2004; Rugg & Wilding, 2000). Interestingly however after incidental encoding, typical familiarity- and recollection-related old/new ERP effects were not apparent in the older adults, even though there were relatively high levels of recognition behaviourally. Both test phases also revealed post-retrieval activity differences across groups, with increased ERP evidence for post-retrieval monitoring in the older adults on Test 1, with both groups showing a similar pattern on the surprise test. This finding suggests that older adults engage in post-retrieval monitoring more widely, maybe as a compensatory mechanism where earlier retrieval processes are diminished. Younger adults may be employing these monitoring processes specifically when the task is more demanding, as seen in the surprise test where items were incidentally encoded and source information was required.

Overall, in this thesis I presented novel behavioural and ERP findings, expanding on the memory-for-foils literature but also contributing to the wider knowledge about how recognition processes change over age. The less conclusive reward modulation studies provide insights into the complexity of using rewards to modulate episodic memory, and I go on in this chapter to provide future directions based on limitations and wider theoretical issues associated with reward effects on memory. The ERP experiment provided more conclusive results, and opens up future directions to continue examining the contribution of item-specific retrieval processes to incidental encoding, and how these are manifest across different age groups.

6.2 Implications of this thesis

Currently there is increasing attention both societally and academically into what is termed 'healthy' cognitive ageing. This is the natural change in cognitive ability as we age, which has interesting implications for our knowledge about abnormal cognitive decline, protective factors against decline, as well as ways or strategies to boost or improve cognition, in particular episodic memory. However, we still have limited and mixed knowledge about healthy ageing with only few studies examining the underlying processes of episodic memory. Experiment 1 and 4 both provide separate yet complementary and novel contributions to knowledge regarding healthy ageing.

Specifically Experiment 1 provides a second look at older adult's ability to strategically constrain retrieval towards a particular source, a very cognitively demanding and effortful task. Here I show that age-related decline in this process is not universal, failing to replicate the one prior study that has examined this issue (Jacoby, Shimizu, Velanova, et al., 2005). As discussed earlier, it is possible that this lack of deficit in the source constrained retrieval effect is due to my sample being non-typical of older adults. They were all University of the Third Age members, very socially active and highly educated. We now know that these attributes are protective factors against cognitive decline (Chan et al., 2018; Josefsson et al., 2012; Nyberg & Pudas, 2019). Nevertheless, even if the Experiment 1 sample is not representative of all older adults, this study provides a novel contribution that showed that many older adults can spontaneously reinstate strategic encoding processes to aid recognition search. This study also showed that many older adults can incidentally encode items during recognition attempts at a matched rate to younger adults in general. This is an important contribution as there is limited research that has examined incidental encoding processes in ageing, with recent suggestions to support the idea that older adults use incidental encoding more often than intentional encoding, especially in their day-to-day activities where they also have more memory complaints (Wagnon et al., 2019).

Not only does Experiment 1 provide knowledge of spontaneous retrieval strategy use in older adults, but older adults were able to outperform younger adults on some tasks. In Experiment 1 were monetary rewards were used, I found that older adults actively tried to not let the external rewards boost their motivation. The older adults as a group showed matched performance on the rewarded and non-rewarded tasks, with many of the older adults not wanting to accept their monetary reward winnings. This behaviour therefore contrasted with younger adults who were very grateful for their winnings, but actually showed a decrease in performance for non-rewarded items.

The undermining effect (Kuhbandner et al., 2016) shows that external rewards can diminish internal motivation, yet here, it seems that in the younger adults rewards do boost internal motivation and strategically focus their encoding strategy on maximising the amount of rewarded items remembered. Older adults in contrast seemed to be more internally motivated to perform their best, regardless of rewards. It is important to consider that rewards, in particular, monetary will not hold the same salience for different groups of people. This is especially potent here as students compared to the older adults recruited with hold the monetary reward with very different esteem, therefore ideally rewards need equally salient to both groups. A suggestion to equalise the salience of rewards offers, which actually is stimulated from the experience I had with the U3A researchers and talking with the older adult participants, could be to offer charitable donations instead of personal monetary winnings. Lots of the older adults' participants left the testing sessions saying they were very pleased they won the money but did not want to receive this fearing it would negatively impact my research budget and many of which then donated this to charity. This may be potent enough to motivate both younger and older participants without undermining their internal motivation to succeed at the task.

Not much has been investigated looking at internal motivation and episodic memory, yet we know that the prospect of and anticipation of high rewards have been seen to increase the likelihood of successful memory encoding (Gruber & Otten, 2010; Gruber et al., 2013) and that rewards increase incidental encoding (Adcock et al., 2006) as well as attentional resource allocation (Anderson, 2016). Looking now more specifically at internal motivation and memory encoding, curiosity has been used previously to signal states of high internal motivation. With interesting findings linking to increased incidental encoding whilst in a state of high curiosity (Gruber et al., 2014; Kang et al., 2009). States of curiosity have also been found to encourage task engagement, attention and semantic encoding even in older adults (McGillivray et al., 2015). Therefore, the wider picture about internal motivation and episodic memory is developing and seems to be more relevant to how older adults engage in tasks with the findings in Experiment 1 supporting this.

Overall, there is limited EEG evidence relating to the neural processes involved in encoding during retrieval attempts, as generally these have been examined with behavioural (Dudukovic et al., 2009; Jacoby, Shimizu, Daniels, et al., 2005; Nairne et al., 2015), fMRI (Buckner et al., 2001; Stark & Okado, 2003), and once previously with the use of EEG oscillations (Vogelsang et al., 2016). These studies have primarily examined if the processes used at encoding are reinstated at recognition. However here I provide novel ERP evidence that the different item-specific retrieval- and decision-related processes that occur during recognition attempts also influence incidental encoding of novel information. Experiment 4 provides evidence in support of distinct neural processes rather than a single "memory strength" process as the basis for recognition. Here I show evidence in support for familiarity, recollection and novelty processing as distinct neural components all involved in recognition attempts and potentially influencing incidental encoding. This is in line with the Dual-Process theory of recognition claiming separate distinct processes for familiarity and recollection (Rugg & Allan, 1995; Wixted, 2007; Woodruff et al., 2006; Yonelinas, 2002). This framework was supported by the ERP findings which showed recollective activity related to hits compared to correct rejections, as well as increased familiarity processing for false alarms compared to increase novelty for fast correct rejections. The separation of highly confident and fast items here also extends on elements of the signal detection theory (SDT, Wixted, 2007) which suggest that highly confident and rapid recognition decisions are further from the old/new decision criterion in terms of their memory strength. A single-process SDT model therefore suggests that rapid and confident new responses are caused by an absence of memory strength, but my findings showed ERP evidence for an additional dimension of novelty strength. The ERP novelty P3 effects seen therefore supports the idea that novelty could be a separate process and not only caused by the lack of familiarity (Daselaar, Fleck, & Cabeza, 2006). Therefore, although potentially novelty and familiarity are two sides of the same coin, the ERP evidence here suggests that these could also be highly correlated but separate processes.

ERP components provide a method of categorising similar patterns of electrical scalp activity based on latency and topography, yet fundamentally components are electrical scalp activity associated with underlying neural processes (Friedman & Johnson, 2000). There are still controversial debates over what functional process components such as the FN400 and FN400 time-windowed effects are actually reflecting (Paller et al., 2007). The main associated processes have been linked to familiarity processing as I suggest here, but semantic priming could also be the underlying mechanism (Voss & Federmeier, 2011). However, from Experiment 4 results it seems that the early time-windowed effects (similar to FN400 or N400 effects) are related to familiarity in this case, given that false alarms elicited a predicted decrease in negativity, whereas correct rejections elicited a more negative going ERP, in line with the FN400 time-windowed effect reflecting a familiarity component. However, it is worth noting that the word stimuli used were obviously semantic in nature and therefore, it could be that foil words that were more semantically related to the studied words became "primed" by the encoding phase, thus eliciting false alarms and a reduced early negativity (Lucas et al., 2012; Woollams et al., 2018). This account was not possible to directly test here, but I go on in the future suggestions section to propose ways of investigating this further. An interesting component not investigated here is the P2 or P200 component evident in early processing (100-300ms). Future studies should investigate this time-window to explore if conceptual priming of items may have contributed to the pattern of results seen here (B. Li, Taylor, Wang, Gao, & Guo, 2017). Yet, even though the separation of the FN400 and the N400 effect is still under debate, and other potential components such as the P200 could have an effect on familiarity, it is worth noting that perhaps these interpretations are interlinked and are manifestations of the same underlying process which changes in scalp location within these early time-windows due to what types of episodic memories are being retrieved (Paller et al., 2007).

Furthermore, given the limited evidence of how memory-related ERPs are manifest in healthy older adults, Experiment 4 provides interesting findings supporting the idea that recognition processes are not universally diminished in older adults or at least not at a uniform rate. In this experiment, after intentional encoding there was evidence for intact familiarity effects, which is in line with some prior studies (Friedman, 2013; Nessler et al., 2007, 2008) and in line with the idea that familiarity processes are spared in healthy ageing (Danckert & Craik, 2013; Fraundorf et al., 2019). Interestingly, however there have been mixed effects linked to neural components in older adults related to familiarity (Duarte et al., 2006; Guillaume et al., 2009; Wang et al., 2012; Wolk et al., 2009). Perhaps there were no early familiarity differences after intentional encoding in Experiment 4 because older adults were very good at the task, with matched discrimination and bias on recognition Test 1.

Yet, there were striking age differences in recognition-related ERPs after incidental encoding on the final surprise test for foils, even though older adults' performance was only slightly lower to that of the younger adults, including source judgments that were very accurate, with even higher accuracy for prior false alarms. This finding raises questions about what the ERP components are actually measuring. If behavioural performance is similar, what does it mean when older and younger adults show neural differences? For example, with older adults there has been a diminished or absent recollective parietal old/new effects (Friedman, 2013; Guillaume et al., 2009; Horne et al., 2020; Wolk et al., 2009), even when intact behavioural measures of recollection are seen (Horne et al., 2020; Nessler et al., 2008). Similarly, in Experiment 4 reduced recollective activity was seen in the older group after intentional encoding even though age group performance was matched, with absent activity after incidental encoding in the older group when performance was slightly lower than in the young group, even though there were correct source judgments indicating recollection of some capacity. These results therefore suggest that either compensatory mechanisms were engaged and enabling older adults to provide correct source responses, or the parietal old/new effect is not the only component indicating recollective experiences, or there is overlapping neural activity on the scalp in older adults that is masking typical parietal old/new effects. This issue brings me back to the methodological problems with using average ERPs, as there may be misleading null effects if there are overlapping ERPs that are opposite in polarity. In the case of older adults, such component overlap "cancelling out" of typical effects has been reported, with a negative going centrally-maximal ERP occurring at the same time as the positive going parietal old/new effect. The negative ERP effect has been argued to occur due to older adults using different retrieval processes, relying on more visual compared to abstract recognition mechanisms (Li et al., 2004; Rugg & Wilding, 2000). From examining the amplitude topographic plots in Experiment 4 there appears to be some left negativity in older adults on both the first and the surprise test, which may be related to the diminished positive activity observed. By no means does this observation provide statistically interpretable evidence for an overlapping component, but it does suggest that diminished or absent ERP components does not rule out the possibility for intact neural processes.

The exploration of novelty P3 time-windowed effects also provide interesting contributions to the literature on novelty processing across age. Seen in Experiment 4 there appears to be a graded effect of novelty strength in younger adults that is enhanced for fast compared to slow correct rejections, whereas in older adults there appears to be a more absolute effect with all correct rejections associated with similar late effect in the P3 time-window. Prior research has shown that older adults show intact novelty responses (Daffner et al., 2005) with others showing decline in P3 amplitude and latency differences (Dinteren et al., 2014). Therefore, perhaps the processes that are underlying novelty P3 effects or effects in this time - window more generally are intact in older age, yet these processes are used in a different way during recognition tasks. Novelty processing was not manipulated in Experiment 4, but the correlational evidence poses an interesting starting point to explore potential differences in novelty processing between young and older adults. I go on to later discuss some potential future directions for this research.

An exciting and practical implication that can be carried forward in future experiments is the collaboration with older adults. The citizen science aspect of Experiment 1 provided an excellent method to firstly gain feedback and guidance on practical aspects of the study, such as font sizes and responses timings. The aim of this feedback was to make the experiment more tailored to older adult's needs, given that older adults often may have worse eye-sight, are harder of hearing, and are generally less familiar with computer devices than younger adults. Secondly, by training older adults as experimenters removed the usual extreme age difference between the experimenter and the participant. This aspect of the collaboration aimed to reduce the stereotype treat potentially felt by older participants (McDaniel et al., 2008), which is known to decrease self-efficacy beliefs in their memory ability (Berry et al., 2010; Fraundorf et al., 2019; West et al., 2009). These positives of our citizen-science collaboration pose an excellent practical implication that can be easily applied in other projects.

Overall, these behavioural and neural findings provide interesting additions to the literature on encoding during retrieval attempts, and the wider healthy ageing literature. I now go on to discuss my reflections on the limitations as well as future directions based on this research.

6.3 Limitations and future directions

There were a range of limitations in the current research, some of which I addressed with paradigm changes throughout the thesis, others that are theoretical where new research came to light as the research progressed, and then there were also practical limitations, for example I was constrained by financial, time or feasibility restrictions. I will reflect on some of the critical limitations and suggest future directions to address these to test the theoretical questions more directly.

6.3.1 External Rewards

The external reward manipulations that I implemented showed limited modulatory effects of episodic memory. Some reward effects on memory were seen in younger adults, but these were only on intentional tasks (in Experiment 1) and did not affect the critical theoretical prediction that rewards might alter recognition processes enough to modulate incidental encoding rates for foils. Some recognition processing changes in response to rewards were seen in Experiment

2 with changes to bias and response times, suggesting that rewards were able to manipulate recognition processes leading to the prioritisation of high value items, suggesting that rewards at retrieval may have the potential to alter incidental encoding rates. Through examining the value-directed remembering literature (Castel et al., 2007, 2011, 2013), which was incorporated in Experiment 2 into the memory-for-foils paradigm (Jacoby, Shimizu, Daniels, et al., 2005), the method of testing incidental encoding of foils could have been more difficult or sensitive. In the typical value-directed remembering study, free-recall tasks are used to test differences in encoding, and these have reliably shown enhancements for high over low value items (Castel et al., 2007, 2011, 2013; Hennessee et al., 2017, 2019; Middlebrooks & Castel, 2018). Therefore, simple alterations to the paradigm designed for Experiment 2 could be made to increase the sensitivity of the final surprise test. Either a free-recall task could be used, which may result in much lower subsequent retrieval rate, but may also show stronger value effects on incidental encoding. The issue of whether external rewards can moderate incidental encoding during retrieval is a complex topic given the nature of how rewards interact with episodic memory processes, and here I have only managed to scrape the surface on this issue. Yet the novel paradigm developed for Experiment 2 could be used as a starting point.

Importantly, prior findings have shown that reward effects on memory tend to be larger when memory is tested after a delay of at least 24-hours, which indicates that reward influences memory consolidation (Murayama & Kitagami, 2014; Spaniol, et al., 2014). The process of consolidation stabilises, reinforces and integrates new important information with the existing long-term memories (Marshall & Born, 2007; Nadel & Moscovitch, 1997). Long term consolidation occurs over a matter of hours and commonly during sleep. Rewards have been seen to interact with consolidation, with rewards potentially functioning as a marker for what information is important to be consolidated, and is thought to enhance consolidation via dopamine release. Consistent with this view, studies often show increased enhancements of memory for highly rewarded items only after a 24-hour delay with fewer effects on memory when tested within the same session (e.g. Adcock et al. 2006; Shohamy & Adcock, 2010; Murayama & Kitagami, 2014; Murayama & Kuhbandner, 2011; see also Patil, Murty, Dunsmoor, Phelps, & Davachi, 2017). Future research should test if rewarded foils are better remembered than non-rewarded foils after a longer delay. Such delays were not used in the current tasks, as there was no scope to examine the effect of consolidation on episodic memory. This is a very vast topic with multiple other interacting factors which would need to be carefully controlled, which was not practical in this series of studies due to time constrains and the additional risk of participant drop-outs. However, examining combined reward and consolidation effects on memory for foils would be a logical follow-on study from Experiment 1, 2 and 3.

Even though there are practical issues regarding how episodic memory retrieval were investigated, as well as theoretical issues around memory consolidation, it is important to note that there are well-documented negative effects of rewards seen in prior literature, which might have been interacting with the expected positive effects of rewards. External rewards do not have the same impact on all individuals, and there are a vast number of other factors that also interact with the salience of rewards. For example, measures of value sensitivity indicate that prioritisation and control processes are engaged less in individuals with explicitly reported lower value sensitivity (Cohen et al., 2017). Rewards can be seen to interact with anxiety and increase performance anxiety and thus hinder memory (Callan & Schweighofer, 2008). This line of evidence supports an ‘undermining effect’, whereby external rewards may diminish and undermine internal motivation (Deci et al., 2001). It is argued that external rewards reduce spontaneous strategy use in intrinsically motivated learners, whereas extrinsically motivated learners show beneficial effects (Kuhbandner et al., 2016) . This research highlights the mixed effects of external rewards on motivation (Deci et al., 2001) with some studies showing no effect on internal motivation (Cameron & Pierce, 1994). These effects are highly relevant to the education literature, but also indicate a complex interplay between reward anticipation in seeking external rewards with internal motivation and other motivationally induced responses. These effects are commonly not controlled for in the memory literature, but is something to consider when looking at subtle processing differences.

One future suggestion could be to measure not only motivation and effort as I did in this thesis, but also measure anxiety levels, and other measures of reward dysfunction. Rewards increase anxiety (Callan & Schweighofer, 2008) but there is also evidence that reward dysfunction reduces both the salience of rewards and the anticipation and motivation to seek rewards (Pizzagalli, Iosifescu, Hallett, Ratner, & Fava, 2008). Clinically this dysfunction is known as anhedonia (Keedwell, Andrew, Williams, Brammer, & Phillips, 2005; Treadway, Buckholtz, Schwartzman, Lambert, & Zald, 2009), but it is also seen in pre-clinical populations (Frey et al., 2015). In the thesis studies, there was an exclusion criterion applied, excluding anyone with a known psychological disorder. However individual differences in reward dysfunction within these ‘healthy’ samples could have been measured to see if this moderated the effect of reward on recognition processes, as large individual differences in memory and reward effects were seen. Here, I would expect those who have signs of reward dysfunction to not attend to the high value items any differently to the low value items, with those that are showing signs of anxiety to potentially show an opposing effect (Callan & Schweighofer, 2008; Murty, LaBar, Hamilton, & Adcock, 2011). This limitation and suggestion highlight the complex nature of interactions between reward and episodic memory, which in the scope of this thesis could not be examined but do pose an interesting and societally relevant line of research that can be continued upon using paradigms designed in this thesis.

6.3.2 Item-specific Recognition Processes

Different retrieval processes that are involved in recognition decisions were seen to alter incidental encoding rates in this thesis. False familiarity was one of two processes leading to increased subsequent recognition, yet it is difficult from these findings to get to the cause or source of the false familiarity. The ERP findings of increased FN400 effects for false alarms coupled with the increased later incorrect source memory of seeing those items during study provides initial insights that there was indeed a false memory for these items, suggesting they were not only associated with incorrect guessing. Yet little can be concluded about if this false memory was caused by extra-experimental familiarity which would vary individually relating to that person's experiences and autobiographical memories, or if the false alarms were caused by within-task factors, such as the semantic relatedness of items. Given that the stimuli were meaningful words and there were words taken from semantically related categories, such as the category 'fruits', then this could be plausible. This design feature is not necessarily a limitation, but there are limits to the interpretation of these findings which would need future examination.

One suggestion would be to try to replicate these effects with less semantically meaningful stimuli. That is, can the ERP and behavioural findings of repeated false familiarity be extrapolated to other stimuli that are semantically impoverished, for example novel face stimuli (Curran & Hancock, 2007), non-words (Otten et al., 2007) or kaleidoscopes (Voss & Paller, 2009). These all have varying degrees of semantic information and the episodic familiarity of these items can be controlled experimentally. Doing so would provide clearer insights into how false episodic familiarity influences encoding during recognition attempts. Using kaleidoscopes, for example, would provide evidence for if non-semantic stimuli produce similar 'familiarity' based neural activation. To my knowledge kaleidoscope images have been used to investigate unconscious memory formation such as through implicit priming (Voss & Paller, 2009), but not directly used with intentional encoding paradigms.

Following on from this suggestion, another related idea could be to increase the FN400 effect by increasing the false memory strength. The DRM paradigm (Roediger & McDermott, 1995) provides a method that allows for the control and manipulation of false memories by increasing the semantic relatedness of items. Potentially this paradigm can be used within the current experimental design to increase the number of false alarms participants make on the first test whilst maintaining a strong memory strength for those items, as the DRM paradigm produces strong false memories of studying items (Gallo & Roediger, 2002; Roediger & McDermott, 1995). This design can be used to study the FN400 effect and examine if inducing false memories via semantic relatedness causes an even larger "carry over" of false familiarity for memory for foils. Increasing the numbers of false alarms is highly useful for ERP analyses, since ERPs can be split into high and low confident false alarms to investigate neural effects associated with memory experience within falsely identified items. For example, highly confident false alarms are likely to be caused by a strong false memory experience that may differ dramatically in brain activity when compared to less confident false alarms, which may not involve a strong memory experience sine they may be more linked to response biases when guessing.

Another suggestion that can be examined by further secondary analysis of the current findings would be to examine how the sematic relatedness of the word stimuli relate to false alarms and memory for foils. This investigation would be difficult to undertake, as either participants would have to rate the semantic relatedness of items, or categories of items would have to be subjectively created. Alternatively, a data driven method could be used where the word lists used at study (which were counterbalanced across participants) are used to predict the likelihood of foil words eliciting false alarms (based on the existing responses to these items). As an example, given that for example list A was shown during encoding, what is the likelihood of the foil word 'APPLE' eliciting a false alarm versus correct rejection response, compared to when word list B was shown during encoding. A classification model would be suitable to classify responses of items given the word lists studied. Such analysis would be complex given that there were multiple word lists rotated to counterbalance stimuli presentation. However, a benefit would be that this analysis could be conducted over multiple experiments within this thesis as they mostly included the same word lists, thereby providing a lot of data.

Alternative suggestions could account for the increase in subsequent recognition of FA items, for example the Compatibility Effect (Craik & Tulving, 1975). In the compatibility effect, words that are responded to with a *'yes'* responses are recognised at a much higher rate later compared to items that received a 'no' response, suggesting that performance can be altered based on prior responses, with yes responses leading to increase latter performance. To overcome this or explore this alternative hypothesis, the questions: *'Is the word old?'* could be counterbalanced with: *'Is the word new?'*. This simply modification would then control for the spread of *yes* and *no* responses to *old* and *new* words. Therefore, it is important to discount alternative hypothesis and conduct further experiments to explore the interaction of false familiarity and encoding processes further.

Similar to familiarity processing, novelty processing here appears to be based on the strength of the novelty processes rather than a threshold, in younger adults at least. Further examination into novelty signals and their effects on subsequent foil memory would provide more insights into the causes of incidental encoding seen in Experiment 4. As these were spontaneously occurring processes, manipulating these were not possible in this study. However, manipulating the novelty of items could be done to test more directly if highly novel foils are associated with fast and highly confident correct rejection judgements, and if these foils are also encoded more frequently compared to less novel foils. The use of 'novel' stimuli may create some issues given that in adults very few stimuli are ever highly novel. Very infrequently used words or even non-words/foreign words could be added as novel targets and these can be intermixed with standard word-lists which vary in familiarity. I would expect to see responses to these novel items to be much more rapid and associated with high confidence in healthy younger adults, supporting the claim that novel items automatically capture attention more due to being distinctive and unique (Krebs et al., 2009; Nyberg, 2005). Manipulating novelty would help determine the underlying mechanisms which detects novelty, and assess if older adults similarly show a graded response to novelty strength or if they use novelty as a threshold during recognition.

6.3.3 Healthy Ageing

The investigation of how healthy ageing affects incidental encoding during recognition provided opportunities to examine spontaneous strategy use, the ability to constrain and control memory search, as well as the underlying neural processes used. The use of strategies in older adults is a developing and expanding field, but this thesis supports the view that older adults can self-initiate strategy use and are able to use cognitive control mechanisms to strategically constrain memory search. Overall, however, less research has investigated the neural processes that are used during recognition in older age.

A limitation of this thesis is in regard to the democratic of the samples used. Primarily the older adult samples were made up of white, middle class, highly educated people, partially due to the recruitment strategy and collaboration with the University of the Third Age, but also partly due to the geographic location of the University of Kent, at Canterbury. Being highly educated and socially active are known protective factors (Chan et al., 2018; Nyberg & Pudas, 2019). Another issue in older adult groups are related to experimenter biases and stereotype threat, in particular about expectations about memory deficits. This can lead to self-selection of older people with good memory, and therefore I tried to reduce likelihood of such biased sampling occurring by advertising the Experiment 1 and 4 as ‘reading’ studies, rather than a memory study. I also avoided informing the participants that there was any comparison to younger participants, as this might have primed the older participants into believing their memory is assumed to be worse. This recruitment strategy may have also selected people who are verbally and cognitively highly performing, yet I felt this was an experimental compromise that was suitable without ethically deceiving participants too much. Wider recruitment is therefore needed to test a more ethnically and socio-economically diverse older group, to see if strategic and neural memory processes are intact in this more representative sample. However, it is important to note that a highly educated and socially active older adult sample is fundamentally more matched to the younger adult samples, which as a standard recruit student. If a more diverse older adults’ group is recruited, then I would also suggest recruiting a more diverse younger adult sample to match performance, education level and social activity more easily.

An unexpected difference seen in the ERP study was of age-differences in frontal slow wave ERPs associated with post-retrieval monitoring. Such monitoring is known to be involved in recognition decisions (Hayama et al., 2008), but the mechanisms underlying these ERP effects are not fully known (Cruse & Wilding, 2009). In Experiment 4, there was an increase in ERPs associated with post-retrieval monitoring in older adults, therefore this raises questions whether this process is compensating or protecting the older adult group from behavioural recognition impairments as seen in prior studies (Fraundorf et al., 2019). Perhaps older adults employ evaluative and monitoring processes to aid their recognition decisions. I could only speculate what effect this might have had on recognition and on incidental encoding.

Suggestions have been made that additional engagement of neural activity in older adults is evidence of a compensatory mechanism (Cabeza et al., 2002; Wolk et al., 2009). Such compensatory mechanisms may have contributed to my ERP findings, however as the younger adults also show ERP markers of post-retrieval monitoring in the surprise recognition test, this effect looks more related to task difficulty. One suggestion could be to directly manipulate task difficulty within-participants. If post-retrieval monitoring is used more when the task is difficult and required more evaluative mechanisms, then I would expect to see an increase in the late frontal right effect, which is related to these processes, on difficult trials as seen in prior literature (Cruse & Wilding, 2009).

Another suggestion for how to better understand post-retrieval monitoring differences across ages could be achieved by altering the task design. I asked participants to provide confidence judgments, which in itself requires introspection and evaluation. From prior literature we know that older adults are less confident overall and less sure about their memory (West et al., 2009), perhaps this led to increases in post-retrieval monitoring, if older adults deliberated about their confidence more than younger adults. It has been shown that activation of the right PFC, an area associated with post-retrieval monitoring and these frontal slow waves seen with ERPs, is activated more for low compared to high confident memories (Henson et al., 2000). Supporting the idea here that increased monitoring related activity is required when memory strengths are weaker, rather than stronger. A suggestion to test this idea would be to remove the confidence judgments in the first recognition test block. If initiating the evaluations of confidence leads to more deliberation and evaluation of memory and this is linked to the increase in post-retrieval monitoring ERP activity, then I would expect to see this ERP effect diminish in the recognition test that did not ask for confidence judgments.

Relatedly, increasing the amount of introspection and evaluation of subjective recognition experience could increase depth of processing of items, and hence incidental encoding. This could be achieved by adding additional subjective measures during the first test in the memory-for-foils paradigm. Additional measures could be added related to if participants remembered contextual details or not in addition to the current confidence measures, or could be linked more to if they relate items to themselves or not. I would expect to see increases in monitoring when these additional subjective ratings are added. This idea is based on the prior source-constrained retrieval literature where increased elaborate engagement with items on the first test led to increased incidental encoding of novel items (Alban, 2013; Alban & Kelley, 2012; Bergström, Vogelsang, Benoit, & Simons, 2015; Danckert et al., 2011; Dudukovic et al., 2009; Gray & Gallo, 2015; Halamish, Goldsmith, & Jacoby, 2012; Jacoby, Shimizu, Daniels, et al., 2005). This would therefore be an interesting addition to Experiment 4, which could provide more knowledge on the neurocognitive processes engaged in most source-constrained retrieval studies that employ strategic and engaging judgment tasks.

A final suggestion, if the sample size in Experiment 4 could be increased would be to split the samples based on performance into high and low recognition performers by a median split within each age group. Some emerging research has shown that in high performing older adults there is intact ERP evidence of recognition processes (Cabeza et al., 2002; de Chastelaine et al., 2016; Duarte et al., 2006; Wang et al., 2016). Here the whole sample performed well on average, but it would be interesting to examine if the ‘low’ performers show more diminished neural activity as compared to the ‘high’ performers. Similarly, it would be important to examine in younger adults whether we can see similar ‘compensatory’ or post-retrieval monitoring ERP effects in the ‘low’ performing younger adults as well.

Overall, the memory-for-foils paradigm and the use of both intentional and incidental encoding tasks provide an interesting paradigm to investigate multiple cognitive processes relating to episodic memory and ageing.

6.4 Impact Statement

Importantly I show here that there is evidence for non-universal cognitive decline over age, with a more holistic view needed when examining memory in older adults both behaviourally and neurally. Currently, we know very little about the brain changes that occur during healthy ageing and what they mean for our understanding of memory processes within young brains. How can we attribute ERP components to brain processes if older adults with similar performance and behavioural ability as younger adults show very different brain activity? We may need to rethink how we understand and interpret ERPs, whilst also examining the potential for overlapping or novel components that might be present in older adults more so than in younger adults. These findings provide a starting point that contributes to the evidence on recognition processes that influence incidental encoding of novel items.

6.5 Conclusions and Final Remarks

The research presented in this thesis examined the impact of healthy ageing, the moderation of external rewards and the influence of item-specific recognition processes on incidental encoding during recognition retrieval attempts. Contrasting to prior findings, my findings show that older adults can spontaneously constrain their recognition processes and outperform younger adults on some recognition measures, suggesting that episodic memory decline in older age, even in cognitively complex and elaborate tasks, is not universal or inevitable. There was little support for external rewards being able to boost or moderate incidental encoding, and simply being more motivated and attending to certain items was not strong enough to alter recognition accuracy beyond altering response biases. I suggested that this effect of rewards does not create a unique enough encoding processing mode that can be reinstated to constrain memory search, as seen in the source-constrained retrieval literature. Strong predictors of incidental encoding however were shown to be linked to more item-elicited recognition processes, with false familiarity and heightened novelty being the best predictors of increased subsequent foil recognition after incidental encoding. The effects of such recognition processes have been overlooked in the incidental encoding literature, with these findings highlighting the contribution of distinct retrieval processes in determining the likelihood that novel information will be later recognised. These findings illustrate the importance of considering sources of false familiarity, as well as the wider implications of incidental encoding processes that occur during recognition attempts. The striking similarity in behavioural performance across ages magnifies the differences in neural activity seen after incidental encoding, providing contributions to our limited understanding of neural processes in older adults, in particular after incidental encoding.

References

Achim, A., & Lepage, M. (2005). Dorsolateral prefrontal cortex involvement in memory post-retrieval monitoring revealed in both item and associative recognition tests. *NeuroImage*, *24*(4), 1113–1121. https://doi.org/10.1016/j.neuroimage.2004.10.036

Adcock, R., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., & Gabrieli, J. (2006). Reward-motivated learning: mesolimbic activation precedes memory formation. *Neuron*, *50*(3), 507–517. https://doi.org/doi.org/10.1016/j.neuron.2006.03.036

Alban, M. (2010). *Controlled Retrieval Processing Among Younger Adults*.

Alban, M. (2013). *Factors Affecting the Likelihood of Constrained Retrieval: The Role of Recollective Details and Cue Distinctiveness*. Retrieved from http://purl.flvc.org/fsu/fd/FSU_migr_etd-7268

Alban, M., & Kelley, C. (2012). Variations in constrained retrieval. *Memory & Cognition*, *40*(5), 681–692. https://doi.org/doi.org/10.3758/s13421-012-0185-5

Ally, B., Waring, J., Beth, E., McKeever, J., Milberg, W., & Budson, A. (2008). Aging memory for pictures: Using high-density event-related potentials to understand the effect of aging on the picture superiority effect. *Neuropsychologia*, *46*(2), 679–689. https://doi.org/10.1016/j.neuropsychologia.2007.09.011

Anderson, B. (2016). The attention habit: How reward learning shapes attentional selection. *Annals of the New York Academy of Sciences*, *1369*(1), 24–39. https://doi.org/10.1111/nyas.12957

Apitz, T., & Bunzeck, N. (2012). Reward modulates the neural dynamics of early visual category processing. *NeuroImage*, *63*(3), 1614–1622. https://doi.org/doi.org/10.1016/j.neuroimage.2012.08.046

Arias-Carrion, O., Stamelou, M., Murillo-Rodriguez, E., Menendez-Gonzalez, M., & Poppel, E. (2010). Dopaminergic reward system: a short integrative review. *International Archives of Medicine*, *3*, 24. https://doi.org/10.1186/1755-7682-3-24

Ariel, R., & Castel, A. (2014). Eyes wide open: Enhanced pupil dilation when selectively studying important information. *Experimental Brain Research*, *232*(1), 337–344. https://doi.org/10.1007/s00221-013-3744-5

Ariel, R., Price, J., & Hertzog, C. (2015). Age-Related Associative Memory Deficits in Value-Based Remembering: The Contribution of Agenda-Based Regulation and Strategy Use. *Psychology and Aging*, *30*(4), 795–808. https://doi.org/10.1037/a0039818

Bäckman, L., Lindenberger, U., Li, S. C., & Nyberg, L. (2010). Linking cognitive aging to alterations in dopamine neurotransmitter functioning: Recent data and future avenues. *Neuroscience and Biobehavioral Reviews*, *34*(5), 670–677. https://doi.org/10.1016/j.neubiorev.2009.12.008

Baddeley, A. (1990a). Retrieval. In *Human Memory: Theory and Practice* (p. 264). Retrieved from https://books.google.co.uk/books?hl=en&lr=&id=fMgm-2NXAXYC&oi=fnd&pg=PP7&dq=human+memory+theory+and+practice+1997&ots=jO2iZMjANl&sig=AbZ_XaM58j55E-TZ9QbiApkFq_A#v=onepage&q=human memory theory and practice 1997&f=false

Baddeley, A. (1990b). When Practice Makes Perfect. In *Human Memory: Theory and Practice* (p. 143). Hove: Lawrence Erlbaum Associates Ltd, Publishers.

Barry, R., Steiner, G., De Blasio, F., Fogarty, J., Karamacoska, D., & MacDonald, B. (2020). Components in the P300: Don't forget the Novelty P3! *Psychophysiology*, *57*(7). https://doi.org/doi.org/10.1111/psyp.13371

Beck, S., Locke, H., Savine, A., Jimura, K., & Braver, T. (2010). Primary and secondary rewards differentially modulate neural activity dynamics during working memory. *PloS One*, *5*(2), e9251. https://doi.org/doi.org/10.1371/journal.pone.0009251

Bergström, Z., Vogelsang, D., Benoit, R., & Simons, J. (2015). Reflections of Oneself: Neurocognitive Evidence for Dissociable Forms of Self-Referential Recollection. *Cerebral Cortex*, *25*(9), 2648–2657. https://doi.org/10.1093/cercor/bhu063

Bergström, Z., Williams, D., Bhula, M., & Sharma, D. (2016). Unintentional and intentional recognition rely on dissociable neurocognitive mechanisms. *Journal of Cognitive Neuroscience*, *28*(11), 1838–1848. https://doi.org/10.1162/jocn_a_01010

Berridge, K., & Robinson, T. (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Research Reviews*, *28*(3), 309–369. https://doi.org/doi.org/10.1016/S0165-0173(98)00019-8

Berridge, K., Robinson, T., & Aldridge, J. (2009). Dissecting components of reward:'liking','wanting', and learning. *Current Opinion in Pharmacology*, *9*(1), 65–73. https://doi.org/doi.org/10.1016/j.coph.2008.12.014

Berry, J., Hastings, E., West, R., Lee, C., & Cavanaugh, J. (2010). Memory Aging: Deficits, Beliefs, and Interventions. In J. Cavanaugh, C. Cavanaugh, J. Berry, & R. West (Eds.), *Praeger perspectives. Aging in America, Vol. 1. Psychological aspects (p. 255–299). Praeger/ABC-CLIO* (Vol. 1, pp. 255–299). Retrieved from http://scholarship.richmond.edu/psychology-faculty-

Berti, S., Vossel, G., & Gamer, M. (2017). The Orienting Response in Healthy Aging: Novelty P3 Indicates No General Decline but Reduced Efficacy for Fast Stimulation Rates. *Frontiers in Psychology*, *8*, 1780. https://doi.org/10.3389/fpsyg.2017.01780

Bijleveld, E., Custers, R., & Aarts, H. (2009). The unconscious eye opener: Pupil dilation reveals strategic recruitment of resources upon presentation of subliminal reward cues. *Psychological Science*, *20*(11), 1313–1315. https://doi.org/10.1111/j.1467-9280.2009.02443.x

Bowen, H., Marchesi, M., & Kensinger, E. (2020). Reward motivation influences response bias on a recognition memory task. *Cognition*, Vol. 203. https://doi.org/10.1016/j.cognition.2020.104337

Brainerd, C., Reyna, V., Wright, R., & Mojardin, A. (2003). Recollection Rejection: False-Memory Editing in Children and Adults. *Psychological Review*, *110*(4), 762–784. https://doi.org/10.1037/0033-295X.110.4.762

Bridger, E., Bader, R., & Mecklinger, A. (2014). More ways than one: ERPs reveal multiple familiarity signals in the word frequency mirror effect. *Neuropsychologia*, *57*(1), 179–190. https://doi.org/10.1016/j.neuropsychologia.2014.03.007

Bröder, A., & Meiser, T. (2007). Measuring source memory. *Journal of Psychology*, *215*(1), 52–60. https://doi.org/10.1027/0044-3409.215.1.52

Buckner, R. (2004). Memory and executive function in aging and ad: Multiple factors that cause decline and reserve factors that compensate. *Neuron*, *44*(1), 195–208. https://doi.org/10.1016/j.neuron.2004.09.006

Buckner, R., Wheeler, M., & Sheridan, M. (2001). Encoding processes during retrieval tasks. *Journal of Cognitive Neuroscience*, *13*(3), 406–415. https://doi.org/doi.org/10.1162/08989290151137430

Butler, A., Fazio, L., & Marsh, E. (2011). The hypercorrection effect persists over a week, but high-confidence errors return. *Psychonomic Bulletin & Review*, *18*(6), 1238–1244. https://doi.org/doi.org/10.3758/s13423-011-0173-y

Butterfield, B., & Metcalfe, J. (2006). The correction of errors committed with high confidence. *Metacognition and Learning*, *1*(1), 69–84. https://doi.org/10.1007/s11409-006-6894-z

Cabeza, Anderson, N., Locantore, J., & McIntosh, A. (2002). Aging gracefully: Compensatory brain activity in high-performing older adults. *NeuroImage*, *17*(3), 1394–1402. https://doi.org/10.1006/nimg.2002.1280

Cabeza, Nyberg, L., & Park, D. (2009). Cognitive Neuroscience of Aging: Linking cognitive and cerebral aging. In R. Cabeza, L. Nyberg, & D. C. Park (Eds.), *Cognitive Neuroscience of Aging: Linking cognitive and cerebral aging*. https://doi.org/10.1093/acprof:oso/9780195156744.001.0001

Callan, D., & Schweighofer, N. (2008). Positive and negative modulation of word learning by reward anticipation. *Human Brain Mapping*, *29*(2), 237–249. https://doi.org/10.1002/hbm.20383

Cameron, J., & Pierce, D. (1994). Reinforcement, Reward, and Intrinsic Motivation: A Meta-Analysis. *Review of Educational Research*, *64*(3), 363–423. https://doi.org/10.3102/00346543064003363

Carlozzi, N. (2011). Shipley Institute of Living Scale. In J. Kreutzer, J. DeLuca, & B. Caplan (Eds.), *Encyclopedia of Clinical Neuropsychology* (pp. 2287–2289). https://doi.org/10.1007/978-0-387-79948-3_1070

Castel, A. (2007). The adaptive and strategic use of memory by older adults: Evaluative processing and value-directed remembering. *Psychology of Learning and Motivation*, *48*, 225–270. https://doi.org/10.1016/S0079-7421(07)48006-9

Castel, A., Benjamin, A., Craik, F., & Watkins, M. (2002). The effects of aging on selectivity and control in short-term recall. *Memory and Cognition*, *30*(7), 1078–1085. https://doi.org/10.3758/BF03194325

Castel, A., Farb, N., & Craik, F. (2007). Memory for general and specific value information in younger and older adults: Measuring the limits of strategic control. *Memory & Cognition*, *35*(4), 689–700. https://doi.org/doi.org/10.3758/BF03193307

Castel, A., Humphreys, K., Lee, S., Galván, A., Balota, D., & McCabe, D. (2011). The development of memory efficiency and value-directed remembering across the life Span: A cross-sectional study of memory and selectivity. *Developmental Psychology*, *47*(6), 1553–1564. https://doi.org/10.1037/a0025623

Castel, A., Murayama, K., Friedman, M., Mcgillivray, S., & Link, I. (2013). Selecting Valuable Information to Remember: Age-Related Differences and Similarities in Self-

Regulated Learning. *Psychology and Aging*, *28*(1), 232–242. https://doi.org/10.1037/a0030678

Chan, D., Shafto, M., Kievit, R., Matthews, F., Spink, M., Valenzuela, M., & Henson, R. (2018). Lifestyle activities in mid-life contribute to cognitive reserve in late-life, independent of education, occupation, and late-life activities. *Neurobiology of Aging*, *70*, 180–183. https://doi.org/10.1016/j.neurobiolaging.2018.06.012

Chiew, K., & Braver, T. (2016). Reward favors the prepared: Incentive and task-informative cues interact to enhance attentional control. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(1), 52–66. https://doi.org/10.1037/xhp0000129

Cohen, M., Cheng, L., Paller, K., & Reber, P. (2019). Separate memory-enhancing effects of reward and strategic encoding. *Journal of Cognitive Neuroscience*, *31*(11), 1658–1673. https://doi.org/10.1162/jocn_a_01438

Cohen, M., Rissman, J., Hovhannisyan, M., Castel, A., & Knowlton, B. (2017). Free recall test experience potentiates strategy-driven effects of value on memory. *Journal of Experimental Psychology: Learning Memory and Cognition*, *43*(10), 1581–1601. https://doi.org/10.1037/xlm0000395

Cohen, M., Rissman, J., Suthana, N., Castel, A., & Knowlton, B. (2014). Value-based modulation of memory encoding involves strategic engagement of fronto-temporal semantic processing regions. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(2), 578–592. https://doi.org/DOI 10.3758/s13415-014-0275-x

Cohen, M., Rissman, J., Suthana, N., Castel, A., & Knowlton, B. (2016). Effects of aging on value-directed modulation of semantic network activity during verbal learning. *NeuroImage*, *125*, 1046–1062. https://doi.org/10.1016/j.neuroimage.2015.07.079

Cohn, M., Emrich, S., & Moscovitch, M. (2008). Age-Related Deficits in Associative Memory: The Influence of Impaired Strategic Retrieval. *Psychology and Aging*, *23*(1), 93–103. https://doi.org/10.1037/0882-7974.23.1.93

Craig, M., Butterworth, K., Nilsson, J., Hamilton, C., Gallagher, P., & Smulders, T. (2016). How does intentionality of encoding affect memory for episodic information? *Learning and Memory*, *23*(11), 648–659. https://doi.org/10.1101/lm.041491.115

Craik, F. (1994). Memory Changes in Aging. *Current Directions in Psychological Science*, *3*(5), 155–158. https://doi.org/10.1007/978-1-4684-6805-2_18

Craik, F. (2002). Levels of processing: Past, present... and future? *Memory*, *10*(5–6), 305–318.

Craik, F., & Lockhart, R. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684.

Craik, F., & McDowd, J. (1987). Age Differences in Recall and Recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(3), 474–479. https://doi.org/10.1037/0278-7393.13.3.474

Craik, F., & Rose, N. (2012). Memory encoding and aging: A neurocognitive perspective. *Neuroscience and Biobehavioral Reviews*, *36*(7), 1729–1739. https://doi.org/10.1016/j.neubiorev.2011.11.007

Craik, F., & Tulving, E. (1975). Depth of processing and the retention of words. *Journal of*

Experimental Psychology, *104*(3), 268–294.

Cruse, D., & Wilding, E. (2009). Prefrontal cortex contributions to episodic retrieval monitoring and evaluation. *Neuropsychologia*, *47*(13), 2779–2789. https://doi.org/10.1016/j.neuropsychologia.2009.06.003

Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition*, *28*(6), 923–938. https://doi.org/doi.org/10.3758/BF03209340

Curran, T., & Cleary, A. (2003). Using ERPs to dissociate recollection from familiarity in picture recognition. *Cognitive Brain Research*, *15*(2), 191–205. https://doi.org/10.1016/S0926-6410(02)00192-1

Curran, T., & Hancock, J. (2007). The FN400 indexes familiarity-based recognition of faces. *NeuroImage*, *36*(2), 464–471. https://doi.org/doi.org/10.1016/j.neuroimage.2006.12.016

Cycowicz, Y., & Friedman, D. (2007). Visual novel stimuli in an ERP novelty oddball paradigm: Effects of familiarity on repetition and recognition memory. *Psychophysiology*, *44*(1), 11–29. https://doi.org/10.1111/j.1469-8986.2006.00481.x

Daffner, K., Ryan, K., Williams, D., Budson, A., Rentz, D., Scinto, L., & Holcomb, P. (2005). Age-related differences in novelty and target processing among cognitively high performing adults. *Neurobiology of Aging*, *26*(9), 1283–1295. https://doi.org/10.1016/j.neurobiolaging.2004.11.007

Danckert, S., & Craik, F. (2013). Does aging affect recall more than recognition memory? *Psychology and Aging*, *28*(4), 902–909. https://doi.org/10.1037/a0033263

Danckert, S., MacLeod, C., & Fernandes, M. (2011). Source-constrained retrieval influences the encoding of new information. *Memory and Cognition*, *39*(8), 1374–1386. https://doi.org/10.3758/s13421-011-0117-9

Daselaar, S., Fleck, M., & Cabeza, R. (2006). Triple dissociation in the medial temporal lobes: recollection, familiarity, and novelty. *Journal of Neurophysiology*, *96*(4), 1902–1911. https://doi.org/01029.2005 [pii]

Daselaar, S., Fleck, M., Dobbins, I., Madden, D., & Cabeza, R. (2006). Effects of healthy aging on hippocampal and rhinal memory functions: An event-related fMRI study. *Cerebral Cortex*, *16*(12), 1771–1782. https://doi.org/10.1093/cercor/bhj112

Dayan, P., & Balleine, B. (2002). Reward, motivation, and reinforcement learning. *Neuron*, *36*(2), 285–298. https://doi.org/10.1016/S0896-6273(02)00963-7

de Chastelaine, M., Mattson, J., Wang, T., Donley, B., & Rugg, M. (2016). The neural correlates of recollection and retrieval monitoring: Relationships with age and recollection performance. *NeuroImage*, *138*, 164–175. https://doi.org/10.1016/j.neuroimage.2016.04.071

Debener, S., Makeig, S., Delorme, A., & Engel, A. (2005). What is novel in the novelty oddball paradigm? Functional significance of the novelty P3 event-related potential as revealed by independent component analysis. *Cognitive Brain Research*, *22*(3), 309–321. https://doi.org/10.1016/j.cogbrainres.2004.09.006

Deci, E. (1975). Cognitive Evaluation Theory: Effects of Extrinsic Rewards on Intrinsic Motivation. *Intrinsic Motivation*, 129–159. https://doi.org/https://doi.org/10.1007/978-1-4613-4446-9_5

Deci, E., Koestner, R., & Richard, R. (2001). Extrinsic rewards and intrinsic motivation in education: Reconsidered once again. *Review of Educational Research*, *71*(1), 1–27. https://doi.org/doi.org/10.3102/00346543071001001

Delgado, M. (2007). Reward-related responses in the human striatum. *Annals of the New York Academy of Sciences*, *1104*, 70–88. https://doi.org/10.1196/annals.1390.002

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

DeLozier, S., & Rhodes, M. (2015). The impact of value-directed remembering on the own-race bias. *Acta Psychologica*, *154*, 62–68. https://doi.org/10.1016/j.actpsy.2014.11.009

DePasque, S., & Tricomi, E. (2015). Effects of intrinsic motivation on feedback processing during learning. *NeuroImage*, *119*, 175–186. https://doi.org/10.1016/j.neuroimage.2015.06.046

Diana, R., Reder, L., Arndt, J., & Park, H. (2006). Models of recognition: A review of arguments in favor of a dual-process account. *Psychonomic Bulletin and Review*, *13*(1), 1–21. https://doi.org/10.3758/BF03193807

Dinteren, R., Arns, M., Jongsma, M., & Kessels, R. (2014). P300 development across the lifespan: A systematic review and meta-analysis. *PLoS ONE*, *9*(2). https://doi.org/10.1371/journal.pone.0087347

Dodson, C., & Schacter, D. (2002). Aging and strategic retrieval processes: Reducing false memories with a distinctiveness heuristic. *Psychology and Aging*, *17*(3), 405. https://doi.org/DOI: 10.1037//0882-7974.17.3.405

Donaldson, W. (1996). The role of decision processes in remembering and knowing. *Memory and Cognition*, *24*(4), 523–533. https://doi.org/10.3758/BF03200940

Duarte, A., Ranganath, C., Trujillo, C., & Knight, R. (2006). Intact recollection memory in high-performing older adults: ERP and behavioral evidence. *Journal of Cognitive Neuroscience*, *18*(1), 33–47. https://doi.org/10.1162/089892906775249988

Dudukovic, N., DuBrow, S., & Wagner, A. (2009). Attention during memory retrieval enhances future remembering. *Memory and Cognition*, *37*(7), 953–961. https://doi.org/10.3758/MC.37.7.953

Duverne, S., Motamedinia, S., & Rugg, M. (2009). Effects of Age on the Neural Correlates of Retrieval Cue Processing are Modulated by Task Demands. *Journal of Cognitive Neuroscience*, *21*(1), 1–17. https://doi.org/10.1162/jocn.2009.21001

Eich, T., Stern, Y., & Metcalfe, J. (2013). The hypercorrection effect in younger and older adults. *Aging, Neuropsychology, and Cognition*, *20*(5), 511–521. https://doi.org/doi.org/10.1080/13825585.2012.754399

Elward, R., Vilberg, K., & Rugg, M. (2015). Motivated memories: Effects of reward and recollection in the core recollection network and beyond. *Cerebral Cortex*, *25*(9), 3159–3166. https://doi.org/10.1093/cercor/bhu109

Eppinger, B., Herbert, M., & Kray, J. (2010). We remember the good things: Age differences in learning and memory. *Neurobiology of Learning and Memory*, *93*(4), 515–521. https://doi.org/doi:10.1016/j.nlm.2010.01.009

Fazio, L., & Marsh, E. (2009). Surprising feedback improves later memory. *Psychonomic Bulletin and Review*, *16*(1), 88–92. https://doi.org/10.3758/PBR.16.1.88

Fazio, L., & Marsh, E. (2010). Correcting false memories. *Psychological Science*, *21*(6), 801–803. https://doi.org/DOI: 10.1177/0956797610371341

Finley, J., & Benjamin, A. (2012). Adaptive and qualitative changes in encoding strategy with experience: Evidence from the test-expectancy paradigm. *Journal of Experimental Psychology: Learning Memory and Cognition*, *38*(3), 632–652. https://doi.org/10.1037/a0026215

Finley, J., Tullis, J., & Benjamin, A. (2010). Metacognitive control of learning and remembering. In *New Science of Learning: Cognition, Computers and Collaboration in Education* (pp. 109–131). https://doi.org/10.1007/978-1-4419-5716-0_6

Finn, B. (2017). A Framework of Episodic Updating: An Account of Memory Updating After Retrieval. *Progress in Brain Research*, *232*, 173–211. https://doi.org/10.1016/bs.plm.2017.03.006

Finnigan, S., Humphreys, M., Dennis, S., & Geffen, G. (2002). ERP ‘old/new’effects: memory strength and decisional factor (s). *Neuropsychologia*, *40*(13), 2288–2304. https://doi.org/doi.org/10.1016/S0028-3932(02)00113-6

Flegal, K., & Lustig, C. (2016). You can go your own way: effectiveness of participant-driven versus experimenter-driven processing strategies in memory training and transfer. *Aging, Neuropsychology, and Cognition*, *23*(4), 389–417. https://doi.org/10.1080/13825585.2015.1108386

Fraundorf, S., Hourihan, K., Peters, R., & Benjamin, A. (2019). Aging and recognition memory: A meta-analysis. *Psychological Bulletin*, *145*(4), 339–371. https://doi.org/10.1037/bul0000185

Frey, A., Malinowska, L., Harley, K., Salhi, L., Iqbal, S., Sharma, S., & McCabe, C. (2015). Investigating subtypes of reward processing deficits as trait markers for depression. *Translational Developmental Psychiatry*, *3*. https://doi.org/doi.org/10.3402/tdp.v3.27517

Friedman, D. (2013). The cognitive aging of episodic memory: A view based on the event-related brain potential. *Frontiers in Behavioral Neuroscience*, *7*(AUG), 111. https://doi.org/10.3389/fnbeh.2013.00111

Friedman, D., Cycowicz, Y., & Gaeta, H. (2001). The novelty P3: an event-related brain potential (ERP) sign of the brain’s evaluation of novelty. *Neuroscience & Biobehavioral Reviews*, *25*(4), 355–373. https://doi.org/10.1182/blood-2009-06-229591

Friedman, D., & Johnson, R. (2000). Event-related potential (ERP) studies of memory encoding and retrieval: A selective review. *Microscopy Research and Technique*, *51*(1), 6–28. https://doi.org/10.1002/1097-0029(20001001)51:1<6::AID-JEMT2>3.0.CO;2-R

Friedman, D., Kazmerski, V., & Cycowicz, Y. (1998). Effects of aging on the novelty P3 during attend and ignore oddball tasks. *Psychophysiology*, *35*(5), 508–520. https://doi.org/10.1017/S0048577298970664

Friedman, D., Nessler, D., & Johnson, R. (2007). Memory encoding and retrieval in the aging brain. *Clinical EEG and Neuroscience*, *38*(1), 1–7. https://doi.org/doi.org/10.1177/155005940703800105

Friedman, D., & Trott, C. (2000). An event-related potential study of encoding in young and older adults. *Neuropsychologia*, *38*(5), 542–557. https://doi.org/10.1016/S0028-3932(99)00122-0

Gallo, D., & Roediger, H. (2002). Variability among word lists in eliciting memory illusions: Evidence for associative activation and monitoring. *Journal of Memory and Language*, *47*(3), 469–497. https://doi.org/10.1016/S0749-596X(02)00013-X

Gardiner, J., & Richardson-Klavehn, A. (2000). Remembering and knowing. In *Learning and Memory: A Comprehensive Reference* (pp. 285–305). https://doi.org/10.1016/B978-012370509-9.00167-4

Geddes, M., Mattfeld, A., Angeles, C., Keshavan, A., & Gabrieli, J. (2018). Human aging reduces the neurobehavioral influence of motivation on episodic memory. *NeuroImage*, *171*, 296–310. https://doi.org/10.1016/j.neuroimage.2017.12.053

Grady, C., McIntosh, A., Rajah, M., Beig, S., & Craik, F. (1999). The effects of age on the neural correlates of episodic encoding. *Cerebral Cortex*, *9*(8), 805–814. https://doi.org/10.1093/cercor/9.8.805

Gray, S. J., & Gallo, D. A. (2015). Disregarding familiarity during recollection attempts: Content-specific recapitulation as a retrieval orientation strategy. *Journal of Experimental Psychology: Learning Memory and Cognition*, *41*(1), 134–147. https://doi.org/10.1037/a0038363

Gross, A., & Rebok, G. (2011). Memory Training and Strategy Use in Older Adults: Results from the ACTIVE Study. *Psychol Aging.*, *26*(3), 503–517. https://doi.org/10.1037/a0022687

Gruber, M., Gelman, B., & Ranganath, C. (2014). States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit. *Neuron*, *84*(2), 486–496. https://doi.org/doi.org/10.1016/j.neuron.2014.08.060

Gruber, M., & Otten, L. (2010). Voluntary control over prestimulus activity related to encoding. *Journal of Neuroscience*, *30*(29), 9793–9800. https://doi.org/10.1523/JNEUROSCI.0915-10.2010

Gruber, M., Watrous, A., Ekstrom, A., Ranganath, C., & Otten, L. (2013). Expected reward modulates encoding-related theta activity before an event. *NeuroImage*, *64*, 68–74. https://doi.org/doi.org/10.1016/j.neuroimage.2012.07.064

Guillaume, C., Clochon, P., Denise, P., Rauchs, G., Guillery-Girard, B., Eustache, F., & Desgranges, B. (2009). Early age-related changes in episodic memory retrieval as revealed by event-related potentials. *NeuroReport*, *20*(2), 191–196. https://doi.org/10.1097/WNR.0b013e32831b44ca

Halamish, V., Goldsmith, M., & Jacoby, L. (2012). Source-constrained recall: Front-end and back-end control of retrieval quality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(1), 1. https://doi.org/10.1037/a0025053

Halsband, T., Ferdinand, N., Bridger, E., & Mecklinger, A. (2012). Monetary rewards influence retrieval orientations. *Cognitive, Affective and Behavioral Neuroscience*, *12*(3), 430–445. https://doi.org/10.3758/s13415-012-0093-y

Han, S., Huettel, S., Raposo, A., Adcock, R., & Dobbins, I. (2010). Functional significance of striatal responses during episodic decisions: Recovery or goal attainment? *Journal of*

Neuroscience, *30*(13), 4767–4775. https://doi.org/10.1523/JNEUROSCI.3077-09.2010

Hauser, T., Eldar, E., Purg, N., Moutoussis, M., & Dolan, R. (2019). Distinct roles of dopamine and noradrenaline in incidental memory. *The Journal of Neuroscience*, 0401–0419. https://doi.org/10.1523/JNEUROSCI.0401-19.2019

Hayama, H., Johnson, J., & Rugg, M. (2008). The relationship between the right frontal old/new ERP effect and post-retrieval monitoring: Specific or non-specific? *Neuropsychologia*, *46*(5), 1211–1223. https://doi.org/10.1016/j.neuropsychologia.2007.11.021

Heit, E., Brockdorff, N., & Lamberts, K. (2004). Strategic processes in false recognition memory. *Psychonomic Bulletin and Review*, *11*(2), 380–386. https://doi.org/10.3758/BF03196586

Hennessee, J., Castel, A., & Knowlton, B. (2017). Recognizing what matters: Value improves recognition by selectively enhancing recollection. *Journal of Memory and Language*, *94*, 195–205. https://doi.org/doi.org/10.1016/j.jml.2016.12.004

Hennessee, J., Patterson, T., Castel, A., & Knowlton, B. (2019). Forget me not: Encoding processes in value-directed remembering. *Journal of Memory and Language*, *106*, 29–39. https://doi.org/10.1016/j.jml.2019.02.001

Henson, R., Campbell, K., Davis, S., Taylor, J., Emery, T., Erzinclioglu, S., … Kievit, R. A. (2016). Multiple determinants of lifespan memory differences. *Scientific Reports*, *6*(1), 1–14. https://doi.org/10.1038/srep32527

Henson, R., & Gagnepain, P. (2010). Predictive, interactive multiple memory systems. *Hippocampus*, *20*(11), 1315–1326. https://doi.org/10.1002/hipo.20857

Henson, R., Rugg, M., Shallice, T., & Dolan, R. (2000). Confidence in recognition memory for words: Dissociating right prefrontal roles in episodic retrieval. *Journal of Cognitive Neuroscience*, *12*(6), 913–923. https://doi.org/10.1162/08989290051137468

Herron, J., Quayle, A., & Rugg, M. (2003). Probability effects on event-related potential correlates of recognition memory. *Cognitive Brain Research*, *16*(1), 66–73. https://doi.org/10.1016/S0926-6410(02)00220-3

Hertzog, C., McGuire, C., Horhota, M., & Jopp, D. (2010). Does believing in “use it or lose it” relate to self-rated memory control, strategy use, and recall? *International Journal of Aging and Human Development*, *70*(1), 61–87. https://doi.org/10.2190/AG.70.1.c

Hidi, S. (2016, March 1). Revisiting the Role of Rewards in Motivation and Learning: Implications of Neuroscientific Research. *Educational Psychology Review*, Vol. 28, pp. 61–93. https://doi.org/10.1007/s10648-015-9307-5

Höltje, G., & Mecklinger, A. (2018). Electrophysiological reward signals predict episodic memory for immediate and delayed positive feedback events. *Brain Research*, *1701*, 64–74. https://doi.org/10.1016/j.brainres.2018.07.011

Horne, E., Koen, J., Hauck, N., & Rugg, M. (2020). Age differences in the neural correlates of the specificity of recollection: An event-related potential study. *Neuropsychologia*, *140*, 107394. https://doi.org/10.1016/j.neuropsychologia.2020.107394

Ingram, K., Mickes, L., & Wixted, J. (2012). Recollection can be weak and familiarity can be strong. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(2),

325. https://doi.org/10.1037/a0025483

Jacoby, L., & Rhodes, M. (2006). False remembering in the aged. *Current Directions in Psychological Science*, *15*(2), 49–53. https://doi.org/10.1111/j.0963-7214.2006.00405.x

Jacoby, L., Shimizu, Y., Daniels, K., & Rhodes, M. (2005). Modes of cognitive control in recognition and source memory: Depth of retrieval. *Psychonomic Bulletin & Review*, *12*(5), 852–857. https://doi.org/doi.org/10.3758/BF03196776

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

Jang, A., Nassar, M., Dillon, D., & Frank, M. (2019). Positive reward prediction errors during decision-making strengthen memory encoding. *Nature Human Behaviour*, *3*(7), 719–732. https://doi.org/10.1038/s41562-019-0597-3

Johnson, J., Hashtroudi, S., & Lindsey, D. (1993). Source Monitoring. *Psychological Bulletin*, *114*(1), 3–28.

Johnson, J., & Rugg, M. (2006). Electrophysiological correlates of retrieval processing: Effects of consistent versus inconsistent retrieval demands. *Journal of Cognitive Neuroscience*, *18*(9), 1531–1544. https://doi.org/10.1162/jocn.2006.18.9.1531

Josefsson, M., Xavier De Luna, †, Pudas, S., Nilsson, L., & Nyberg, L. (2012). Genetic and Lifestyle Predictors of 15-Year Longitudinal Change in Episodic Memory. *Journal of the American Geriatrics Societ*, *60*(12), 2308–2312. https://doi.org/10.1111/jgs.12000

Jou, J., & Flores, S. (2013). How are false memories distinguishable from true memories in the Deese-Roediger-McDermott paradigm? A review of the findings. *Psychological Research*, Vol. 77, pp. 671–686. https://doi.org/10.1007/s00426-012-0472-6

Kafkas, A., & Montaldi, D. (2014). Two separate, but interacting, neural systems for familiarity and novelty detection: A dual-route mechanism. *Hippocampus*, *24*(5), 516–527.

Kang, M., Hsu, M., Krajbich, I., Loewenstein, G., McClure, S., Wang, J., & Camerer, C. (2009). The wick in the candle of learning: Epistemic curiosity activates reward circuitry and enhances memory. *Psychological Science*, *20*(8), 963–973. https://doi.org/10.1111/j.1467-9280.2009.02402.x

Kang, M., & Pashler, H. (2014). Is the benefit of retrieval practice modulated by motivation? *Journal of Applied Research in Memory and Cognition*, *3*(3), 183–188. https://doi.org/10.1016/j.jarmac.2014.05.006

Kantner, J., & Lindsay, S. (2013). Top-down constraint on recognition memory. *Memory & Cognition*, *41*(3), 465–479. https://doi.org/10.3758/s13421-012-0265-6

Karpicke, J., Lehman, M., & Aue, W. (2014). Retrieval-Based Learning. An Episodic Context Account. In *Psychology of Learning and Motivation - Advances in Research and Theory* (Vol. 61, pp. 237–284). https://doi.org/10.1016/B978-0-12-800283-4.00007-1

Keedwell, P., Andrew, C., Williams, S., Brammer, M., & Phillips, M. (2005). The neural correlates of anhedonia in major depressive disorder. *Biological Psychiatry*, *58*(11), 843–853. https://doi.org/doi:10.1016/j.biopsych.2005.05.019

Kinsbourne, M., & George, J. (1974). The mechanism of the word-frequency effect on recognition memory. *Journal of Verbal Learning and Verbal Behavior*, *13*(1), 63–69. https://doi.org/10.1016/S0022-5371(74)80031-9

Kirchhoff, B. (2009). Individual differences in episodic memory: The role of self-initiated encoding strategies. *Neuroscientist*, *15*(2), 166–179. https://doi.org/10.1177/1073858408329507

Kirchhoff, B., Anderson, B., Barch, D., Jacoby, L., & Kirchhoff, B. (2012). Cognitive and Neural Effects of Semantic Encoding Strategy Training in Older Adults. *NeuroImage*, *62*(3), 1956–1964. https://doi.org/10.1093/cercor/bhr129

Kizilirmak, J., Thuerich, H., Folta-Schoofs, K., Schott, B., & Richardson-Klavehn, A. (2016). Neural correlates of learning from induced insight: A case for reward-based episodic encoding. *Frontiers in Psychology*, *7*(1693). https://doi.org/10.3389/fpsyg.2016.01693

Knight, R. (1996). Contribution of human hippocampal region to novelty detection. *Nature*, *383*(6597), 256–259. https://doi.org/10.1038/383256a0

Koen, J., & Yonelinas, A. (2016). Recollection, not familiarity, decreases in healthy ageing: Converging evidence from four estimation methods. *Memory*, *24*(1), 75–88. https://doi.org/10.1080/09658211.2014.985590

Koriat, A., & Goldsmith, M. (1996). Monitoring and Control Processes in the Strategic Regulation of Memory Accuracy. *Psychological Review*, *103*(3), 490–517. https://doi.org/10.1037/0033-295X.103.3.490

Krans, J., Näring, G., Becker, E., & Holmes, E. (2009). Intrusive trauma memory: A review and functional analysis. *Applied Cognitive Psychology*, Vol. 23, pp. 1076–1088. https://doi.org/10.1002/acp.1611

Krebs, R., Schott, B., Schütze, H., & Düzel, E. (2009). The novelty exploration bonus and its attentional modulation. *Neuropsychologia*, *47*(11), 2272–2281. https://doi.org/10.1016/j.neuropsychologia.2009.01.015

Krendl, A., Ambady, N., & Kensinger, E. (2015). The dissociable effects of stereotype threat on older adults' memory encoding and retrieval. *Journal of Applied Research in Memory and Cognition*, *4*(2), 103–109. https://doi.org/10.1016/j.jarmac.2015.02.001

Kuhbandner, C., Aslan, A., Emmerdinger, K., & Murayama, K. (2016). Providing Extrinsic Reward for Test Performance Undermines Long-Term Memory Acquisition. *Frontiers in Psychology*, *7*, 79. https://doi.org/10.3389/fpsyg.2016.00079

Lamont, R., Swift, H., & Abrams, D. (2015). A Review and Meta-Analysis of Age-Based Stereotype Threat: Negative Stereotypes, Not Facts, Do the Damage. *Psychology and Aging*, *30*(1), 180. https://doi.org/10.1037/a0038586

Lang, P., Bradley, M., & Cuthbert, B. (2005). IAPS: Affective ratings of pictures and instruction manual. *Emotion*.

Li, J., Morcom, A., & Rugg, M. (2004). The effects of age on the neural correlates of successful episodic retrieval: an ERP study. *Cognitive, Affective, & Behavioral Neuroscience*, *4*(3), 279–293. https://doi.org/10.3758/CABN.4.3.279

Li, Taylor, J. R., Wang, W., Gao, C., & Guo, C. (2017). Electrophysiological signals associated with fluency of different levels of processing reveal multiple contributions to recognition memory. *Consciousness and Cognition*, *53*(April), 1–13. https://doi.org/10.1016/j.concog.2017.05.001

Lisman, J., & Grace, A. (2005, June 2). The hippocampal-VTA loop: Controlling the entry of information into long-term memory. *Neuron*, Vol. 46, pp. 703–713. https://doi.org/10.1016/j.neuron.2005.05.002

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

Lucas, H., Taylor, J., Henson, R., & Paller, K. (2012). Many roads lead to recognition: Electrophysiological correlates of familiarity derived from short-term masked repetition priming. *Neuropsychologia*, *50*(13), 3041–3052. https://doi.org/10.1016/j.neuropsychologia.2012.07.036

Luck, S. J. (2014). An Introduction to the Second Edition Event-Related Potential Technique. In *MIT Press* (Second, Vol. 33). https://doi.org/10.1073/pnas.0703993104

Luo, L., & Craik, F. (2008). Aging and memory: A cognitive approach. *Canadian Journal of Psychiatry*, *53*(6), 346–353. https://doi.org/10.1177/070674370805300603

Madan, C., & Spetch, M. (2012). Is the enhancement of memory due to reward driven by value or salience? *Acta Psychologica*, *139*(2), 343–349. https://doi.org/10.1016/j.actpsy.2011.12.010

Malmberg, K. (2008). Recognition memory: A review of the critical findings and an integrated theory for relating them. *Cognitive Psychology*, *57*(4), 335–384. https://doi.org/10.1016/j.cogpsych.2008.02.004

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, *87*(3), 252–271. https://doi.org/10.1037/0033-295X.87.3.252

Mandler, G., Goodman, G., & Wilkes-Gibbs, D. (1982). The word-frequency paradox in recognition. *Memory & Cognition*, *10*(1), 33–42. https://doi.org/10.3758/BF03197623

Marini, F., Marzi, T., & Viggiano, M. (2011). "Wanted!" The effects of reward on face recognition: electrophysiological correlates. *Cognitive, Affective, & Behavioral Neuroscience*, *11*(4), 627–643. https://doi.org/DOI 10.3758/s13415-011-0057-7

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of Neuroscience Methods*, *164*, 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Marsh, R., Meeks, J., Cook, G., Clark-Foos, A., Hicks, J., & Brewer, G. (2009). Retrieval constraints on the front end create differences in recollection on a subsequent test. *Journal of Memory and Language*, *61*(3), 470–479. https://doi.org/10.1016/j.jml.2009.06.005

Marshall, L., & Born, J. (2007). The contribution of sleep to hippocampus-dependent memory consolidation. *Trends in Cognitive Sciences*, Vol. 11, pp. 442–450. https://doi.org/10.1016/j.tics.2007.09.001

Mason, A., Farrell, S., Howard-Jones, P., & Ludwig, C. (2017). The role of reward and reward uncertainty in episodic memory. *Journal of Memory and Language*, *96*, 62–77. https://doi.org/doi.org/10.1016/j.jml.2017.05.003

Mather, M., & Schoeke, A. (2011). Positive outcomes enhance incidental learning for both younger and older adults. *Frontiers in Neuroscience*, *5*, 129. https://doi.org/10.3389/fnins.2011.00129 [doi]

May, C., Rahhal, T., Berry, E., & Leighton, E. (2005). Aging, source memory, and emotion. *Psychology and Aging*, *20*(4), 571–578. https://doi.org/10.1037/0882-7974.20.4.571

McDaniel, M. A., Einstein, G. O., & Jacoby, L. L. (2008). New considerations in aging and memory. *The Handbook of Aging and Cognition*, *4*, 251–310.

Mcdermott, K. B., & Roediger, H. L. (1998). *Attempting to Avoid Illusory Memories: Robust False Recognition of Associates Persists under Conditions of Explicit Warnings and Immediate Testing.*

McDonough, I. M., Bui, D. C., Friedman, M. C., & Castel, A. D. (2015). Retrieval monitoring is influenced by information value: The interplay between importance and confidence on false memory. *Acta Psychologica*, *161*, 7–17. https://doi.org/10.1016/j.actpsy.2015.07.017

McGillivray, S., Murayama, K., & Castel, A. (2015). Thirst for Knowledge: The Effects of Curiosity and Interest on Memory in Younger and Older Adults. *Psychology and Aging*, *30*(4), 835–841. https://doi.org/10.1037/a0039801

Mecklinger, A. (2010). The control of long-term memory: Brain systems and cognitive processes. *Neuroscience and Biobehavioral Reviews*, *34*(7), 1055–1065. https://doi.org/10.1016/j.neubiorev.2009.11.020

Mecklinger, A., & Jäger, T. (2012). Episodic memory storage and retrieval: Insights from electrophysiological measures. In *Neuroimaging of Human Memory: Linking Cognitive Processes to Neural Systems*. https://doi.org/10.1093/acprof:oso/9780199217298.003.0020

Metcalfe, J. (2017). Learning from Errors. *Annual Review of Psychology*, *68*(1), 465–489. https://doi.org/10.1146/annurev-psych-010416-044022

Metcalfe, J., & Finn, B. (2011). People's hypercorrection of high-confidence errors: Did they know it all along? *Journal of Experimental Psychology: Learning Memory and Cognition*, *37*(2), 437–448. https://doi.org/10.1037/a0021962

Middlebrooks, C. (2018). *Encoding and retrieval influences on the strategic study of important information* (UCLA). Retrieved from https://cloudfront.escholarship.org/dist/prd/content/qt8m76k0z9/qt8m76k0z9.pdf

Middlebrooks, C., & Castel, A. (2018). Self-regulated learning of important information under sequential and simultaneous encoding conditions. *Journal of Experimental Psychology: Learning Memory and Cognition*, *44*(5), 779–792. https://doi.org/10.1037/xlm0000480

Middlebrooks, C., Murayama, K., & Castel, A. (2016). The value in rushing: Memory and selectivity when short on time. *Acta Psychologica*, *170*, 1–9. https://doi.org/doi.org/10.1016/j.actpsy.2016.06.001

Miendlarzewska, E., Bavelier, D., & Schwartz, S. (2016). Influence of reward motivation on human declarative memory. *Neuroscience and Biobehavioral Reviews*, Vol. 61, pp. 156–176. https://doi.org/10.1016/j.neubiorev.2015.11.015

Migo, E., Mayes, A., & Montaldi, D. (2012). Measuring recollection and familiarity: Improving the remember/know procedure. *Consciousness and Cognition*, *21*(3), 1435–1455. https://doi.org/doi.org/10.1016/j.concog.2012.04.014

Mitchell, K., & Johnson, M. (2000). Source Monitoring: Attributing mental experiences. In E. Tulving & F.I.M. Craik (Eds.) (Ed.), *The Oxford handbook of memory* (pp. 179–195). New York: Oxford University Press.

Mognon, A., Jovicich, J., Bruzzone, L., & Buiatti, M. (2011). ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology*, *48*(2), 229–240. https://doi.org/10.1111/j.1469-8986.2010.01061.x

Morcom, A. (2016). Mind Over Memory: Cuing the Aging Brain. *Current Directions in Psychological Science*, *25*(3), 143–150. https://doi.org/10.1177/0963721416645536

Morcom, A., Good, C., Frackowiak, S., & Rugg, M. (2003). Age effects on the neural correlates of successful memory encoding. *Brain*, *126*(1), 213–229. https://doi.org/10.1093/brain/awg020

Morcom, A., Li, J., & Rugg, M. (2007). Age effects on the neural correlates of episodic retrieval: Increased cortical recruitment with matched performance. *Cerebral Cortex*, *17*(11), 2491–2506. https://doi.org/10.1093/cercor/bhl155

Morcom, A., & Rugg, M. (2004). Effects of age on retrieval cue processing as revealed by ERPs. *Neuropsychologia*, *42*(11), 1525–1542. https://doi.org/10.1016/j.neuropsychologia.2004.03.009

Mullet, H., & Marsh, E. (2016). Correcting false memories: Errors must be noticed and replaced. *Memory and Cognition*, *44*(3), 403–412. https://doi.org/10.3758/s13421-015-0571-x

Murayama, K., & Kitagami, S. (2014). Consolidation power of extrinsic rewards: reward cues enhance long-term memory for irrelevant past events. *Journal of Experimental Psychology: General*, *143*(1), 15. https://doi.org/http://dx.doi.org/10.1037/a0031992

Murayama, K., & Kuhbandner, C. (2011). Money enhances memory consolidation–But only for boring material. *Cognition*, *119*(1), 120–124. https://doi.org/doi:10.1016/j.cognition.2011.01.001

Murty, V., & Adcock, R. (2014). Enriched encoding: reward motivation organizes cortical networks for hippocampal detection of unexpected events. *Cerebral Cortex (New York, N.Y.: 1991)*, *24*(8), 2160–2168. https://doi.org/10.1093/cercor/bht063 [doi]

Murty, V., LaBar, K., Hamilton, D., & Adcock, R. (2011). Is all motivation good for learning? Dissociable influences of approach and avoidance motivation in declarative memory. *Learning and Memory*, *18*(11), 712–717. https://doi.org/10.1101/lm.023549.111

Nadel, L., & Moscovitch, M. (1997). Memory consolidation, retrograde amnesia, and the temporal lobe. *Current Opinion in Neurobiology*, *7*(2), 217–227. https://doi.org/doi.org/10.1016/S0959-4388(97)80010-4

Nairne, J., Pandeirada, J., VanArsdall, J., & Blunt, J. (2015). Source-constrained retrieval and survival processing. *Memory & Cognition*, *43*(1), 1–13. https://doi.org/10.3758/s13421-014-0456-4

Nairne, J., Thompson, S., & Pandeirada, J. (2007). Adaptive memory: Survival processing enhances retention. *Journal of Experimental Psychology: Learning Memory and Cognition*, *33*(2), 263–273. https://doi.org/10.1037/0278-7393.33.2.263

Nasreddine, Z., Phillips, N., 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. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Nessler, D., Friedman, D., Johnson, R., & Bersick, M. (2007). Does repetition engender the same retrieval processes in young and older adults? *NeuroReport*, *18*(17), 1837–1840. https://doi.org/10.1097/WNR.0b013e3282f16d9f

Nessler, D., Johnson, R., Bersick, M., & Friedman, D. (2008). Age-related ERP differences at retrieval persist despite age-invariant performance and left-frontal negativity during encoding. *Neuroscience Letters*, *432*, 151–156. https://doi.org/10.1016/j.neulet.2007.12.016

Nyberg, L. (2005, August). Any novelty in hippocampal formation and memory? *Current Opinion in Neurology*, Vol. 18, pp. 424–428. https://doi.org/10.1097/01.wco.0000168080.99730.1c

Nyberg, L., Habib, R., McIntosh, A., & Tulving, E. (2000). Reactivation of encoding-related brain activity during memory retrieval. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(20), 11120–11124. https://doi.org/97/20/11120 [pii]

Nyberg, L., Lövdén, M., Riklund, K., Lindenberger, U., & Bäckman, L. (2012). Memory aging and brain maintenance. *Trends in Cognitive Sciences*, Vol. 16, pp. 292–305. https://doi.org/10.1016/j.tics.2012.04.005

Nyberg, L., & Pudas, S. (2019). Successful Memory Aging. *Annu. Rev. Psychol*, *70*, 219–243. https://doi.org/10.1146/annurev-psych-010418

Old, S., & Naveh-Benjamin, M. (2008). Differential Effects of Age on Item and Associative Measures of Memory: A Meta-Analysis. *Psychology and Aging*, *23*(1), 104–118. https://doi.org/10.1037/0882-7974.23.1.104

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, *2011*. https://doi.org/10.1155/2011/156869

Otmakhova, N., Duzel, E., Deutch, A., & Lisman, J. (2013). The hippocampal-VTA loop: the role of novelty and motivation in controlling the entry of information into long-term memory. In *Intrinsically motivated learning in natural and artificial systems* (pp. 235–254). https://doi.org/DOI 10.1007/978-3-642-32375-1 10

Otten, L., & Rugg, M. (2001). Electrophysiological correlates of memory encoding are task-dependent. *Cognitive Brain Research*, *12*(1), 11–18. https://doi.org/10.1016/S0926-6410(01)00015-5

Otten, L., & Rugg, M. (2005). Interpreting Event-Related Brain Potentials. In *Event-related potentials: A methods handbook* (pp. 3–16).

Otten, L., Sveen, J., & Quayle, A. (2007). Distinct patterns of neural activity during memory formation of nonwords versus words. *Journal of Cognitive Neuroscience*, *19*(11), 1776–1789. https://doi.org/10.1162/jocn.2007.19.11.1776

Paller, K., Voss, J., & Boehm, S. (2007). Validating neural correlates of familiarity. *Trends in Cognitive Sciences*, *11*(6), 243–250. https://doi.org/doi:10.1016/j.tics.2007.04.002

Pashler, H., Cepeda, N., Wixted, J., & Rohrer, D. (2005). When does feedback facilitate learning of words? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(1), 3. https://doi.org/https://doi.org/10.1037/0278-7393.31.1.3

Patil, A., Murty, V. P., Dunsmoor, J. E., Phelps, E. A., & Davachi, L. (2017). Reward retroactively enhances memory consolidation for related items. *Learning & Memory*, *24*(1), 65–69. https://doi.org/10.1101/lm.042978

Pause, B., Zlomuzica, A., Kinugawa, K., Mariani, J., Pietrowsky, R., & Dere, E. (2013). Perspectives on episodic-like and episodic memory. *Frontiers in Behavioral Neuroscience*, *7*, 33. https://doi.org/https://doi.org/10.3389/fnbeh.2013.00033

Paxton, J., Barch, D., Storandt, M., & Braver, T. (2006). Effects of environmental support and strategy training on older adults' use of context. *Psychology and Aging*, *21*(3), 499–509. https://doi.org/10.1037/0882-7974.21.3.499

Peirce, J. (2007). PsychoPy-Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Peirce, J., Gray, J., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., … Lindeløv, J. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Picton, Bentin, S., Berg, P., Donchin, E., Hillyard, S., Johnson, R., … Taylor, M. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology*, *37*(02), 127–152. https://doi.org/https://doi.org/10.1111/1469-8986.3720127

Pizzagalli, D., Iosifescu, D., Hallett, L., Ratner, K., & Fava, M. (2008). Reduced hedonic capacity in major depressive disorder: Evidence from a probabilistic reward task. *Journal of Psychiatric Research*, *43*(1), 76–87. https://doi.org/10.1016/j.jpsychires.2008.03.001

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, Vol. 118, pp. 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Pyc, M., & Rawson, K. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? *Journal of Memory and Language*, *60*(4), 437–447. https://doi.org/https://doi.org/10.1016/j.jml.2009.01.004

Ranganath, C., & Rainer, G. (2003). Cognitive neuroscience: Neural mechanisms for detecting and remembering novel events. *Nature Reviews Neuroscience*, *4*(3), 193–202. https://doi.org/10.1038/nrn1052

Ratcliff, R., & Starns, J. (2013). Modeling confidence judgments, response times, and multiple choices in decision making: recognition memory and motion discrimination. *Psychological Review*, *120*(3), 697–719. https://doi.org/10.1037/a0033152

Rhodes, M., Witherby, A., Castel, A., & Murayama, K. (2017). Explaining the forgetting bias effect on value judgments: The influence of memory for a past test. *Memory and*

Cognition, *45*(3), 362–374. https://doi.org/10.3758/s13421-016-0674-z

Richardson-Klavehn, A., & Bjork, R. (1988). Measures Of Memory. *Annual Review of Psychology*, *39*(1), 475–543. https://doi.org/10.1146/annurev.psych.39.1.475

Richardson, C., Bucks, R., & Hogan, A. (2011). Effects of aging on habituation to novelty: An ERP study. *International Journal of Psychophysiology*, *79*(2), 97–105. https://doi.org/10.1016/j.ijpsycho.2010.09.007

Robinson, L., Stevens, L., Threapleton, C., Vainiute, J., McAllister-Williams, H., & Gallagher, P. (2012). Effects of intrinsic and extrinsic motivation on attention and memory. *Acta Psychologica*, *141*(2), 243–249. https://doi.org/10.1016/j.actpsy.2012.05.012

Robison, M., & Unsworth, N. (2017). Working memory capacity, strategic allocation of study time, and value-directed remembering. *Journal of Memory and Language*, *93*, 231–244. https://doi.org/10.1016/j.jml.2016.10.007

Roediger, H., & Karpicke, J. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*(3), 249–255. https://doi.org/https://doi.org/10.1111/j.1467-9280.2006.01693.x

Roediger, H., & McDermott, K. (1995). Creating False Memories: Remembering Words Not Presented in Lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 803–814. https://doi.org/10.1037/0278-7393.21.4.803

Rouhani, N., Norman, K., & Niv, Y. (2018). Dissociable effects of surprising rewards on learning and memory. *Journal of Experimental Psychology: Learning Memory and Cognition*, *44*(9), 1430–1443. https://doi.org/10.1037/xlm0000518

Rugg, M., & Allan, K. (1995). Event-related potential studies of human memory. *The Cognitive Neurosciences*, 789–801. Retrieved from https://www.researchgate.net/publication/32889391

Rugg, M., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, *11*(6), 251–257. https://doi.org/doi:10.1016/j.tics.2007.04.004

Rugg, M., & Morcom, A. (2009). The Relationship Between Brain Activity, Cognitive Performance, and Aging: The Case of Memory. In *Cognitive Neuroscience of Aging: Linking cognitive and cerebral aging* (pp. 132–154). https://doi.org/10.1093/acprof:oso/9780195156744.003.0006

Rugg, M., Otten, L., & Henson, R. (2002). The neural basis of episodic memory: Evidence from functional neuroimaging. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *357*(1424), 1097–1110. https://doi.org/10.1098/rstb.2002.1102

Rugg, M., Walla, P., Schloerscheidt, A., Fletcher, P., Frith, C., & Dolan, R. (1998). Neural correlates of depth of processing effects on recollection: Evidence from brain potentials and positron emission tomography. *Experimental Brain Research*, *123*(1–2), 18–23. https://doi.org/10.1007/s002210050540

Rugg, M., & Wilding, E. (2000). Retrieval processing and episodic memory. *Trends in Cognitive Sciences*, Vol. 4, pp. 108–115. https://doi.org/10.1016/S1364-6613(00)01445-5

Saczynski, J., Rebok, G., Whitfield, K., & Plude, D. (2007). Spontaneous production and use of mnemonic strategies in older adults. *Experimental Aging Research*, *33*(3), 273–294. https://doi.org/10.1080/03610730701318899

Sakaki, M., Yagi, A., & Murayama, K. (2018). Curiosity in old age: A possible key to achieving adaptive aging. *Neuroscience and Biobehavioral Reviews*, Vol. 88, pp. 106–116. https://doi.org/10.1016/j.neubiorev.2018.03.007

Salthouse, T. (2000). Aging and measures of processing speed. *Biological Psychology*, *54*(1–3), 35–54. https://doi.org/10.1016/S0301-0511(00)00052-1

Schacter, D. (1999). The seven sins of memory: Insights from psychology and cognitive neuroscience. *American Psychologist*, *54*(3), 182–203. https://doi.org/10.1037/0003-066X.54.3.182

Seymour, B., Singer, T., & Dolan, R. (2007). The neurobiology of punishment. *Nature Reviews Neuroscience*, *8*(4), 300–311. https://doi.org/10.1038/nrn2119

Shafto, M., Henson, R., Matthews, F., Taylor, J., Emery, T., Erzinclioglu, S., … Tyler, L. (2019). Cognitive Diversity in a Healthy Aging Cohort: Cross-Domain Cognition in the Cam-CAN Project. *Journal of Aging and Health*. https://doi.org/10.1177/0898264319878095

Shigemune, Y., Tsukiura, T., Kambara, T., & Kawashima, R. (2014). Remembering with gains and losses: effects of monetary reward and punishment on successful encoding activation of source memories. *Cerebral Cortex*, *24*(5), 1319–1331. https://doi.org/10.1093/cercor/bhs415 [doi]

Shigemune, Y., Tsukiura, T., Nouchi, R., Kambara, T., & Kawashima, R. (2017). Neural mechanisms underlying the reward-related enhancement of motivation when remembering episodic memories with high difficulty. *Human Brain Mapping*, *38*(7), 3428–3443. https://doi.org/10.1002/hbm.23599

Shimizu, Y., & Jacoby, L. (2005). Similarity-guided depth of retrieval: Constraining at the front end. *Canadian Journal of Experimental Psychology*, *59*(1), 17–21. https://doi.org/doi.org/10.1037/h0087455

Shohamy, D., & Adcock, R. (2010). Dopamine and adaptive memory. *Trends in Cognitive Sciences*, *14*(10), 464–472. https://doi.org/doi:10.1016/j.tics.2010.08.002

Snodgrass, J., & Corwin, J. (1988). Pragmatics of Measuring Recognition Memory: Applications to Dementia and Amnesia. In *Journal of Experimental Psychology: General* (Vol. 117). https://doi.org/10.1037/0096-3445.117.1.34

Spaniol, J., Bowen, H., Wegier, P., & Grady, C. (2015). Neural responses to monetary incentives in younger and older adults. *Brain Research*, *1612*, 70–82. https://doi.org/10.1016/j.brainres.2014.09.063

Spaniol, J., Schain, C., & Bowen, H. (2014). Reward-enhanced memory in younger and older adults. *The Journals of Gerontology.Series B, Psychological Sciences and Social Sciences*, *69*(5), 730–740. https://doi.org/10.1093/geronb/gbt044 [doi]

Spencer, W., & Raz, N. (1995). Differential Effects of Aging on Memory for Content and Context: A Meta-Analysis. *Psychology and Aging*, *10*(4), 527–539. https://doi.org/10.1037/0882-7974.10.4.527

Squire, L. (2004). Memory systems of the brain: A brief history and current perspective.

Neurobiology of Learning and Memory, *82*(3), 171–177. https://doi.org/10.1016/j.nlm.2004.06.005

Squire, L., Knowlton, B., & Musen, G. (1993). The Structure and Organization of Memory. *Annual Review of Psychology*, *44*(1), 453–495. https://doi.org/10.1146/annurev.ps.44.020193.002321

Stark, C., & Okado, Y. (2003). Making memories without trying: medial temporal lobe activity associated with incidental memory formation during recognition. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *23*(17), 6748–6753. https://doi.org/23/17/6748 [pii]

Suengas, A., Gallego-Largo, T., & Simón, T. (2010). Age-related changes in recognition and response criterion. *Spanish Journal of Psychology*, *13*(2), 557–571. https://doi.org/10.1017/S1138741600002249

Taconnat, L., Clarys, D., Vanneste, S., Bouazzaoui, B., & Isingrini, M. (2007). Aging and strategic retrieval in a cued-recall test: The role of executive functions and fluid intelligence. *Brain and Cognition*, *64*(1), 1–6. https://doi.org/10.1016/j.bandc.2006.09.011

Taylor, J., & Henson, R. (2012). Could masked conceptual primes increase recollection? The subtleties of measuring recollection and familiarity in recognition memory. *Neuropsychologia*, *50*(13), 3027–3040. https://doi.org/10.1016/j.neuropsychologia.2012.07.029

Téllez-Alanís, B., & Cansino, S. (2004). Incidental and intentional encoding in young and elderly adults. *Neuroreport*, *15*(11), 1819–1823. https://doi.org/DOI:10.1097/01.wnr.0000137075.41257.98

Treadway, M., Buckholtz, J., Schwartzman, A., Lambert, W., & Zald, D. (2009). Worth the 'EEfRT'? The effort expenditure for rewards task as an objective measure of motivation and anhedonia. *PloS One*, *4*(8), e6598. https://doi.org/doi:10.1371/journal.pone.0006598

Tulving, E. (1993). What is episodic memory? *Current Directions in Psychological Science*, *2*(3), 67–70. https://doi.org/https://doi.org/10.1111/1467-8721.ep10770899

Tulving, E., Markowitsch, H., Craik, F., Habib, R., & Houle, S. (1996). Novelty and Familiarity Activations in PET Studies of Memory Encoding and Retrieval. *Cerebral Cortex*, *6*(1), 71–79. https://doi.org/10.1093/cercor/6.1.71

Tun, P., Wingfield, A., Rosen, M., & Blanchard, L. (1998). Response latencies for false memories: Gist-based processes in normal aging. *Psychology and Aging*, *13*(2), 230–241. https://doi.org/10.1037/0882-7974.13.2.230

Underwood, B. (1978). Recognition memory as a function of length of study list. *Bulletin of the Psychonomic Society*, *12*(2), 89–91. https://doi.org/10.3758/BF03329636

Vilberg, K., Moosavi, R., & Rugg, M. (2006). The relationship between electrophysiological correlates of recollection and amount of information retrieved. *Brain Research*, *1122*(1), 161–170. https://doi.org/https://doi.org/10.1016/j.brainres.2006.09.023

Vogelsang, D., Bonnici, H., Bergström, Z., Ranganath, C., & Simons, J. (2016). Goal-directed mechanisms that constrain retrieval predict subsequent memory for new "foil" information. *Neuropsychologia*, *89*, 356–363. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2016.07.016

Vogelsang, D., Gruber, M., Bergström, Z., Ranganath, C., & Simons, J. (2018). Alpha oscillations during incidental encoding predict subsequent memory for new "foil" information. *Journal of Cognitive Neuroscience*, *30*(Early Access), 667–679. https://doi.org/10.1162/jocn_a_01234

Voss, J., & Federmeier, K. (2011). FN400 potentials are functionally identical to N400 potentials and reflect semantic processing during recognition testing. *Psychophysiology*, *48*(4), 532–546. https://doi.org/10.1111/j.1469-8986.2010.01085.x

Voss, J., & Paller, K. (2009). An electrophysiological signature of unconscious recognition memory. *Nature Neuroscience*, *12*(3), 349–355. https://doi.org/10.1038/nn.2260

Wagnon, C., Wehrmann, K., Klöppel, S., & Peter, J. (2019). Incidental learning: A systematic review of its effect on episodic memory performance in older age. *Frontiers in Aging Neuroscience*, *10*(JUL), 173. https://doi.org/10.3389/fnagi.2019.00173

Wang, T., de Chastelaine, M., Minton, B., & Rugg, M. (2012). Effects of age on the neural correlates of familiarity as indexed by ERPs. *Journal of Cognitive Neuroscience*, *24*(5), 1055–1068. https://doi.org/10.1162/jocn_a_00129

Wang, T., Johnson, J., De Chastelaine, M., Donley, B., & Rugg, M. (2016). The Effects of Age on the Neural Correlates of Recollection Success, Recollection-Related Cortical Reinstatement, and Post-Retrieval Monitoring. *Cerebral Cortex*, *26*(4), 1698–1714. https://doi.org/10.1093/cercor/bhu333

Wegesin, D., Friedman, D., Varughese, N., & Stern, Y. (2002). Age-related changes in source memory retrieval: an ERP replication and extension. In *Cognitive Brain Research* (Vol. 13). https://doi.org/https://doi.org/10.1016/S0926-6410(01)00126-4

Weidemann, C., & Kahana, M. (2016). Assessing recognition memory using confidence ratings and response times. *Royal Society Open Science*, *3*(4). https://doi.org/10.1098/rsos.150670

West, R., Dark-Freudeman, A., & Bagwell, D. (2009). Goals-feedback conditions and episodic memory: Mechanisms for memory gains in older and younger adults. *Memory*, *17*(2), 233–244. https://doi.org/10.1080/09658210802236407

Wilding, E. (2000). In what way does the parietal ERP old/new effect index recollection? *International Journal of Psychophysiology*, *35*(1), 81–87. https://doi.org/10.1016/S0167-8760(99)00095-1

Wilding, E., & Nobre, A. (2001). Task-switching and memory retrieval processing: Electrophysiological evidence. *Cognitive Neuroscience & Neuropsychology*, *12*(16), 3613–3617. https://doi.org/DOI: 10.1097/00001756-200111160-00048

Wilding, E., & Rugg, M. (1996). An event-related potential study of recognition memory with and without retrieval of source. *Brain*, *119*(3), 889–905. https://doi.org/10.1093/brain/119.3.889

Wilson, M. (1988). MRC Psycholinguistic Database: Machine-usable dictionary. Version 2. *Behavior Research Methods, Instruments, & Computers*, *20*(1), 6–10. https://doi.org/https://doi.org/10.3758/BF03202594

Wimmer, G., Braun, E., Daw, N., & Shohamy, D. (2014). Episodic memory encoding interferes with reward learning and decreases striatal prediction errors. *Journal of Neuroscience*, *34*(45), 14901–14912. https://doi.org/10.1523/JNEUROSCI.0204-

14.2014

Wittmann, B., Bunzeck, N., Dolan, R., & Düzel, E. (2007). Anticipation of novelty recruits reward system and hippocampus while promoting recollection. *NeuroImage*, *38*(1), 194–202. https://doi.org/doi:10.1016/j.neuroimage.2007.06.038

Wittmann, B., Dolan, R., & Duzel, E. (2011). Behavioral specifications of reward-associated long-term memory enhancement in humans. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *18*(5), 296–300. https://doi.org/10.1101/lm.1996811 [doi]

Wittmann, B., Schott, B., Guderian, S., Frey, J., Heinze, H., & Düzel, E. (2005). Reward-related FMRI activation of dopaminergic midbrain is associated with enhanced hippocampus-dependent long-term memory formation. *Neuron*, *45*(3), 459–467. https://doi.org/DOI 10.1016/j.neuron.2005.01.010 Reward-Related

Wixted, J. (2007). Dual-Process Theory and Signal-Detection Theory of Recognition Memory. *Psychological Review*, *114*(1), 152. https://doi.org/10.1037/0033-295X.114.1.152

Wixted, J., & Mickes, L. (2010). A Continuous Dual-Process Model of Remember/Know Judgments. *Psychological Review*, *117*(4), 1025–1054. https://doi.org/10.1037/a0020874

Wolk, D., Schacter, D., Lygizos, M., Sen, N. M., Holcomb, P., Daffner, K., & Budson, A. (2006). ERP correlates of recognition memory: Effects of retention interval and false alarms. *Brain Research*, *1096*(1), 148–162. https://doi.org/10.1016/j.brainres.2006.04.050

Wolk, D., Sen, N., Chong, H., Riis, J., McGinnis, S., Holcomb, P., & Daffner, K. (2009). ERP correlates of item recognition memory: Effects of age and performance. *Brain Research*, *1250*, 218–231. https://doi.org/10.1016/j.brainres.2008.11.014

Wolosin, S., Zeithamova, D., & Preston, A. (2012). Reward modulation of hippocampal subfield activation during successful associative encoding and retrieval. *Journal of Cognitive Neuroscience*, *24*(7), 1532–1547. https://doi.org/https://doi.org/10.1162/jocn_a_00237

Woodruff, C., Hayama, H., & Rugg, M. (2006). Electrophysiological dissociation of the neural correlates of recollection and familiarity. *Brain Research*, *1100*(1), 125–135. https://doi.org/10.1016/j.brainres.2006.05.019

Woollams, A., Taylor, J., Karayanidis, F., & Henson, R. (2018). ERPs associated with masked priming of test cues reveal multiple potential contributions to recognition memory. *Journal of Cognitive Neuroscience*, *62*, 1–2. https://doi.org/10.16107/j.cnki.mmte.2018.0891

Yan, C., Li, Y., Zhang, Q., & Cui, L. (2018). Monetary incentives at retrieval promote recognition of involuntarily learned emotional information. *NeuroReport*, *29*(4), 259–265. https://doi.org/10.1097/WNR.0000000000000932

Yonelinas, A. (2001). Consciousness, control, and confidence: The 3 Cs of recognition memory. *Journal of Experimental Psychology: General*, *130*(3), 361–379. https://doi.org/10.1037/0096-3445.130.3.361

Yonelinas, A. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*(3), 441–517.

https://doi.org/10.1006/jmla.2002.2864

Yonelinas, Aly, M., Wang, W. C., & Koen, J. D. (2010). Recollection and familiarity: Examining controversial assumptions and new directions. *Hippocampus*, *20*(11), 1178–1194. https://doi.org/10.1002/hipo.20864

Zawadzka, K., Hanczakowski, M., & Wilding, E. (2017). Late consequences of early selection: When memory monitoring backfires. *Journal of Memory and Language*, *92*, 114–127. https://doi.org/10.1016/j.jml.2016.06.003

Appendix

Appendix 1: Chapter 2- Experiment 1

Appendix 1.1: Chapter 2- Experiment 1 – Experiment Paperwork

Appendix 1.1.1: Experiment 1 – Study information sheet

Study Information Sheet

Title of Project:	Word processing.	**Ethics Approval Number:**	201614805240054138
Investigator(s):	Louisa Salhi Dr Zara Bergstrom	**Researcher Email:**	lmas6@kent.ac.uk

You are being invited to take part in a research study. Before you decide, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully. Please ask if anything is unclear, and do not feel rushed into making a decision.

Study Aim

This is a research study to examine how people learn and remember based on different amounts of reward, in the form of money.

Eligibility Requirements

To participate in this study you must be over the age of 60 and have English as your native language. You must also have normal or corrected to normal vision (i.e. normal with glasses or contact lenses). Participants must not be Dyslexic. Participants need to be neurologically healthy and not currently taking any psychoactive medication such as anti-depressants.

What will happen to me in the study?

You will perform a task that involves judging words as well as later memorising words that will appear on a computer screen. Before each task you will be told what to do and about what rewards you can win. After the study you will receive your winnings. The maximum amount you can win is £5.00. You will be given full instructions and get an opportunity to practice some elements of the task before the real experiment begins. After the study you may also be asked to complete some short questionnaires and the Montreal Cognitive Assessment. The complete session will take around 50 minutes.

What will happen if I don't want to carry on with the study?

We want to emphasise that you may withdraw from the study at any time, and if you refuse to take part or decide to withdraw you will not suffer any penalty or loss of rights. If you do choose to withdraw or are no longer able to participate, then unless verbally directed otherwise the study investigators will keep the data collected up to that point. Your participation can be withdrawn by the study investigators.

Are there any negative side-effects/risks?

We do not anticipate any significant side-effects although some participants may feel a little tired after the test session.

What will happen to the results of the research study?

The results of this study will be used to better understand the mental processes that support word and/or picture memory and how they are affected by rewards. The results of the study may be published for scientific purposes, but your records or identity will not be revealed unless required by law.

What happens to the information I provide?

Participation in this study guarantees confidentiality of the information you provide. No one apart from members of the research team and carefully selected research collaborators will have any access to the information you provide. Your name and any other identifying information will be stored separately from your data in a securely locked filing cabinet. Questionnaires will be stored in a securely locked room for as long as is required by the Data Protection Act, and then they will be destroyed by our confidential shredding service. You may withdraw from the study at any time, and if you refuse to take part or if you decide to withdraw, you will not suffer any penalty or loss of rights. Your participation in the study can be withdrawn by the study investigators.

What if there is a problem?

If you have a concern about any aspect of the study then you should speak with Dr Zara Bergström, who is director of the study. She can be reached on 01227 827507. If you remain unhappy and wish to complain formally, you can do this through the School of Psychology Chair of Ethics. Further details can be obtained from the School of Psychology General Office on 01227 824775.

Who is organising the research?

The research is organised by Louisa Salhi, a PhD student, under the supervisor of Dr Zara Bergström who works within the School of Psychology.

Who has reviewed this study?

The study has been approved by the School of Psychology, University of Kent Research Ethics committee.

If you would like a copy of the consent form or information sheet to keep, please ask the researcher. If you have any complaints or concerns about this research, you can direct these, in writing, to the Chair of the Psychology Research Ethics Committee by email at: psychethics@kent.ac.uk. Alternatively, you can contact us by post at: Ethics Committee Chair, School of Psychology, University of Kent, Canterbury, CT2 7NP.

Appendix 1.1.2: Experiment 1 - Consent Form

RESEARCH INFORMED CONSENT FORM

Title of Project:	Word processing.	**Ethics Approval Number:**	201614805240054138
Investigator(s):	Louisa Salhi Dr Zara Bergström	**Researcher Email:**	lmas6@kent.ac.uk

Please read the following statements and, if you agree, initial the corresponding box to confirm agreement:

Signatures:

	Initials
I confirm that I have read and understand the information sheet for the above study. I have had the opportunity to consider the information, ask questions about the risks involved and have had these answered satisfactorily.	________
I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason.	________
I understand that my personal information and data will be treated as strictly confidential and will not be made public or shared with any other person outside of the research team.	________
I understand that my data may be published in a research article, but that no personal details will be divulged and that it will not be possible to identify any responses as my own.	
I confirm that I am over 18 years of age and that I freely agree to participate in this study.	

______________________	______________________	______________________
Name of participant (block capitals)	Date	Signature
______________________	______________________	______________________
Researcher (block capitals)	Date	Signature

If you would like a copy of this consent form to keep, please ask the researcher. If you have any complaints or concerns about this research, you can direct these, in writing, to the Chair of the Psychology Research Ethics Committee by email at: psychethics@kent.ac.uk. Alternatively, you can contact us by post at: Ethics Committee Chair, School of Psychology, University of Kent, Canterbury, CT2 7NP.

Appendix 1.1.3: Experiment 1 - Instructions – Read aloud by Instructors
These vary depending on the group

- **Firstly in this study you will be asked to judge words in specific ways. All instructions about this will be on the screen and verbally told to you.**

  You will be later tested on these words however if you follow our instructions and only do the task instructed this will help you on the test. This task is difficult so don't worry if you feel that the task is hard.

Pleasantness Judgments Task

For the next few trials you will see words presented on the screen.

Please can you make judgments about the pleasantness of the words in this list and respond by using the keyboard. You have a few seconds to make this judgment.

Do you think the word is pleasant?

YES the word is pleasant press the green button

NO the word is not pleasant press the red button

This is meant to be subjective and a gut feeling and just focus on the decision and don't do anything else.

Try to answer as quickly and accurately as you can. It is easier if you keep your fingers on the keyboard responses.

If you have any questions please ask the experimenter, if not press 'space bar' to continue.

Visual search task

During the next few trials you will see pictures on the screen, and we want you to look for a white triangle. Not all the images will have a white triangle, but if you see one respond with the green key to say Yes there is a white triangle.

Is there a WHITE TRIANGLE on the screen?

Use the Keyboard to respond by pressing:

YES there is a WHITE TRIANGLE : press the green button

NO there is not a WHITE TRIANGLE : press the red button

Try to be as quick and accurate as possible.

If you have any questions please ask the experimenter, if not press 'space bar' to continue.

Vowel Judgment Task

For the next few trials you will see words presented on the screen.

Please can you make judgments about the vowels in each word in this list and respond by using the keyboard. You have a few seconds to make this judgment.

Is there an 'O' or 'U' in the word?

YES there is an 'O' or 'U' press the green button

NO there is not an 'O' or 'U' press the red button

Try to answer as quickly and accurately as you can. It is easier if you keep your fingers on the keyboard responses.

If you have any questions please ask the experimenter, if not press 'space bar' to continue.

Visual search task

Is there a WHITE TRIANGLE on the screen?

Use the Keyboard to respond by pressing:

YES there is a WHITE TRIANGLE : press the green button

NO there is not a WHITE TRIANGLE : press the red button

Try to be as quick and accurate as possible.

Pleasantness Recognition Test

For the next few trials you will see words presented on the screen, some of these words are new and some you have seen before. The **old words in this recognition task are from the pleasantness judgment task.** The task is self-paced but try to answer swiftly.

Try to recognize which words were in the pleasantness judgment task and which words are new. Try to be as quick and accurate as possible. Use the keyboard to respond:

Do you recognise the word?

YES it's an OLD word press the green button

NO it's a NEW word press the red button

If you have any questions please ask the experimenter, if not press 'space bar' to continue.

Visual search task

Is there a WHITE TRIANGLE on the screen?

Use the Keyboard to respond by pressing:

YES there is a WHITE TRIANGLE : press the green button

NO there is not a WHITE TRIANGLE : press the red button

Try to be as quick and accurate as possible.

Vowel Judgment Recognition Test

For the next few trials you will see words presented on the screen, some of these words are new and some you have seen before. The **old words in this recognition task are from the vowel judgment task where you judged if the words had the letter ‘O’ or ‘U’ in them all the words from that task are included regardless of how you judged them previously.**

The task is self-paced but try to answer swiftly.

During the next few trials try to recognize which words were in the vowel counting task and which words are new. Try to be as quick and accurate as possible. Use the keyboard to respond:

Do you recognise the word?

YES it’s an OLD word press the green button

NO it’s a NEW word press the red button

If you have any questions please ask the experimenter, if not press ‘space bar’ to continue.

Visual search task

Is there a WHITE TRIANGLE on the screen?

Use the Keyboard to respond by pressing:

YES there is a WHITE TRIANGLE : press the green button

NO there is not a WHITE TRIANGLE : press the red button

Try to be as quick and accurate as possible.

Rewarding Memory Task

For the next few trials you will see words presented on the screen, you will not have to respond to these words but please try to memorize the list of words. Each word will be on the screen for a limited amount of time.

For every correctly recognised word you will get 10p for correctly identifying an old or new word.

Overall for this memory task you can win up to £5.00!!!

If you have any questions please ask the experimenter, if not press ‘space bar’ to continue.

Visual search task

Is there a WHITE TRIANGLE on the screen?

Use the Keyboard to respond by pressing:

YES there is a WHITE TRIANGLE : press the green button

NO there is not a WHITE TRIANGLE : press the red button

Try to be as quick and accurate as possible.

Memory Task

For the next few trials you will see words presented on the screen, you will not have to respond to these words but please try to memorize the list of words.

Each word will be on the screen for a limited amount of time.

If you have any questions please ask the experimenter, if not press ‘space bar’ to continue.

Visual search task

Is there a WHITE TRIANGLE on the screen?

Use the Keyboard to respond by pressing:

YES there is a WHITE TRIANGLE : press the green button

NO there is not a WHITE TRIANGLE : press the red button

Try to be as quick and accurate as possible.

Rewarding Recognition task

For the next few trials you will see words presented on the screen, some of these words are new and some you have seen before.

During the next few trials try to recognize which words were in **rewarding memory** task and which words are new. Try to be as quick and accurate as possible.

For every correctly recognised word you will get 10p for correctly identifying an old or new word.

Overall for this memory task you can win up to £5.00!!!

Use the keyboard to respond:

Do you recognise the word?

YES it's an OLD word press the green button

NO it's a NEW word press the red button

If you have any questions please ask the experimenter, if not press 'space bar' to continue.

Visual search task

Is there a WHITE TRIANGLE on the screen?

Use the Keyboard to respond by pressing:

YES there is a WHITE TRIANGLE : press the green button

NO there is not a WHITE TRIANGLE : press the red button

Try to be as quick and accurate as possible.

Recognition task

For the next few trials you will see words presented on the screen, some of these words are new and some you have seen before.

During the next few trials try to recognize which words were in **the memory** task and which words are new. Try to be as quickly and accurately as possible.

Use the keyboard to respond:

Do you recognise the word?

YES it's an OLD word press the green button

NO it's a NEW word press the red button

If you have any questions please ask the experimenter, if not press 'space bar' to continue.

Visual search task

Is there a WHITE TRIANGLE on the screen?

Use the Keyboard to respond by pressing:

YES there is a WHITE TRIANGLE : press the green button

NO there is not a WHITE TRIANGLE : press the red button

Try to be as quick and accurate as possible.

Final Surprise Test

Now you are about to start the final test.

For the next few trials you will see words presented on the screen. Some of these words are new and some you have seen before during the study.

If you recognise any words from the **whole study** respond yes.

This includes all study and test phases so it can contains any words that have been presented on the screen so far!

Do you recognise the word?

Use the Keyboard to respond by pressing:

YES it's an OLD word press the green button

NO it's a NEW word press the red button

If you have any questions please ask the experimenter.

When you are ready to start press SPACE BAR.

Appendix 1.1.4: Experiment 1 – Post Study Questionnaire

Post Study Questionnaire

PARTICPANT ID:_____DATE: / ______ /

1: Did you anticipate the **final surprise test** which was at the end? *(Please circle)*

YES NO

2: Did you use any techniques to help you learn the words in the **rewarding memory learning task**? If so, please briefly tell us about these techniques. *(Please circle and write in the space below)*

YES NO

____

____

2: Did you use any techniques to help you learn the words in the **standard memory task** (when you were just asked to remember words but were not given any rewards)?

If so, please briefly tell us about these techniques. *(Please circle and write in the space below)*

YES NO

____

____

4: Can you rate how motivated you were to win the maximum amount of money during the **rewarding memory learning task**:

1. *Unmotivated 5- indifferent 10- very motivated*

1 2 3 4 5 6 7 8 9 10

5: Can you rate how much effort you put into learning the words during the **rewarding memory learning task**:

1. *no effort 5- indifferent 10- high effort*

1 2 3 4 5 6 7 8 9 10

6: Can you rate how motivated you were to win the maximum amount of money during the **rewarding memory Test**:

1. *unmotivated 5- indifferent 10- very motivated*

1 2 3 4 5 6 7 8 9 10

7: Can you rate how much effort you put into recognising the words during the **rewarding memory Test**:

1. *no effort 5- indifferent 10- high effort*

1 2 3 4 5 6 7 8 9 10

8: Can you rate how much effort you put into learning the words during the **standard memory Test**:

2. *no effort 5- indifferent 10- high effort*

1 2 3 4 5 6 7 8 9 10

9: Can you rate how much effort you put into recognising the words during the **standard memory Test**:

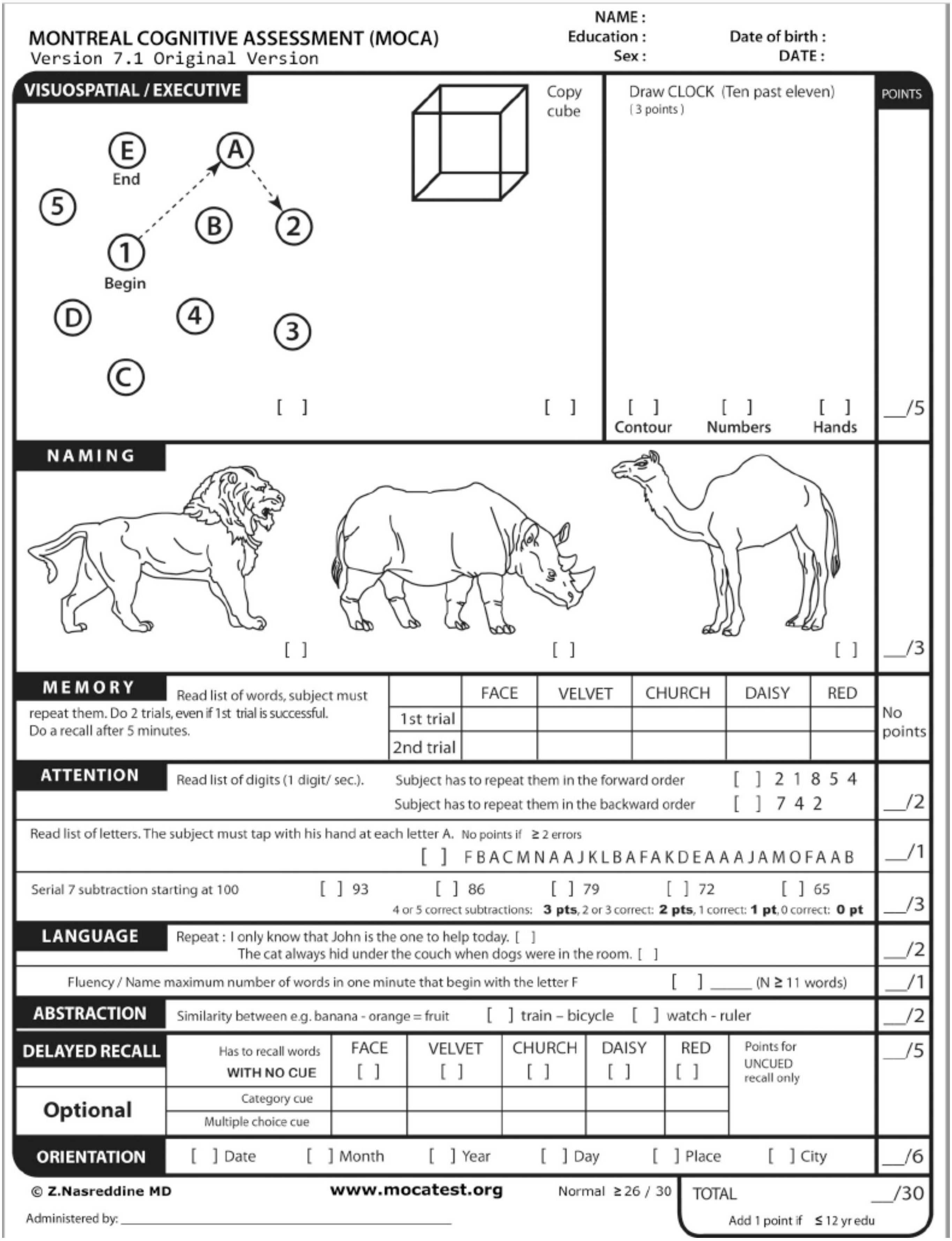
1- no effort 5- indifferent 10- high effort

1 2 3 4 5 6 7 8 9 10

10: Please write down how many years you have been in education.

Do you have any other **comments/feedback about the study**?

Appendix 1.1.5: Montreal cognitive Assessment and Instructions



MONTREAL COGNITIVE ASSESSMENT (MOCA)
Version 7.1 Original Version

NAME :
Education :
Date of birth :
Sex :
DATE :

		POINTS
VISUOSPATIAL / EXECUTIVE [ ] Copy cube [ ]	Draw CLOCK (Ten past eleven) (3 points) [ ] Contour [ ] Numbers [ ] Hands	__/5
NAMING [ ] [ ] [ ]		__/3

MEMORY Read list of words, subject must repeat them. Do 2 trials, even if 1st trial is successful. Do a recall after 5 minutes.

	FACE	VELVET	CHURCH	DAISY	RED	
1st trial						No points
2nd trial						

ATTENTION Read list of digits (1 digit/ sec.).
Subject has to repeat them in the forward order [ ] 2 1 8 5 4
Subject has to repeat them in the backward order [ ] 7 4 2 __/2

Read list of letters. The subject must tap with his hand at each letter A. No points if ≥ 2 errors
[ ] F B A C M N A A J K L B A F A K D E A A A J A M O F A A B __/1

Serial 7 subtraction starting at 100 [ ] 93 [ ] 86 [ ] 79 [ ] 72 [ ] 65
4 or 5 correct subtractions: **3 pts**, 2 or 3 correct: **2 pts**, 1 correct: **1 pt**, 0 correct: **0 pt** __/3

LANGUAGE Repeat : I only know that John is the one to help today. [ ]
The cat always hid under the couch when dogs were in the room. [ ] __/2

Fluency / Name maximum number of words in one minute that begin with the letter F [ ] ____ (N ≥ 11 words) __/1

ABSTRACTION Similarity between e.g. banana - orange = fruit [ ] train – bicycle [ ] watch - ruler __/2

DELAYED RECALL		FACE	VELVET	CHURCH	DAISY	RED		
	Has to recall words **WITH NO CUE**	[ ]	[ ]	[ ]	[ ]	[ ]	Points for UNCUED recall only	__/5
Optional	Category cue							
	Multiple choice cue							

ORIENTATION [ ] Date [ ] Month [ ] Year [ ] Day [ ] Place [ ] City __/6

© Z.Nasreddine MD **www.mocatest.org** Normal ≥ 26 / 30 TOTAL ___/30
Add 1 point if ≤ 12 yr edu

Administered by: ______________________________

Montreal Cognitive Assessment
(MoCA)

Administration and Scoring Instructions

The Montreal Cognitive Assessment (MoCA) was designed as a rapid screening instrument for mild cognitive dysfunction. It assesses different cognitive domains: attention and concentration, executive functions, memory, language, visuoconstructional skills, conceptual thinking, calculations, and orientation. Time to administer the MoCA is approximately 10 minutes. The total possible score is 30 points; a score of 26 or above is considered normal.

1. Alternating Trail Making:

Administration: The examiner instructs the subject: *"Please draw a line, going from a number to a letter in ascending order. Begin here* [point to (1)] *and draw a line from 1 then to A then to 2 and so on. End here* [point to (E)]."

Scoring: Allocate one point if the subject successfully draws the following pattern:
1 –A- 2- B- 3- C- 4- D- 5- E, without drawing any lines that cross. Any error that is not immediately self-corrected earns a score of 0.

2. Visuoconstructional Skills (Cube):

Administration: The examiner gives the following instructions, pointing to the **cube**: *"Copy this drawing as accurately as you can, in the space below"*.

Scoring: One point is allocated for a correctly executed drawing.

- Drawing must be three-dimensional
- All lines are drawn
- No line is added
- Lines are relatively parallel and their length is similar (rectangular prisms are accepted)

A point is not assigned if any of the above-criteria are not met.

3. Visuoconstructional Skills (Clock):

Administration: Indicate the right third of the space and give the following instructions: *"Draw a* ***clock****. Put in all the numbers and set the time to 10 past 11"*.

Scoring: One point is allocated for each of the following three criteria:

- Contour (1 pt.): the clock face must be a circle with only minor distortion acceptable (e.g., slight imperfection on closing the circle);
- Numbers (1 pt.): all clock numbers must be present with no additional numbers; numbers must be in the correct order and placed in the approximate quadrants on the clock face; Roman numerals are acceptable; numbers can be placed outside the circle contour;
- Hands (1 pt.): there must be two hands jointly indicating the correct time; the hour hand must be clearly shorter than the minute hand; hands must be centred within the clock face with their junction close to the clock centre.

A point is not assigned for a given element if any of the above-criteria are not met.

4. **Naming**:

Administration: Beginning on the left, point to each figure and say: *"Tell me the name of this animal".*

Scoring: One point each is given for the following responses: (1) lion (2) rhinoceros or rhino (3) camel or dromedary.

5. **Memory**:

Administration: The examiner reads a list of 5 words at a rate of one per second, giving the following instructions: *"This is a memory test. I am going to read a list of words that you will have to remember now and later on. Listen carefully. When I am through, tell me as many words as you can remember. It doesn't matter in what order you say them".* Mark a check in the allocated space for each word the subject produces on this first trial. When the subject indicates that (s)he has finished (has recalled all words), or can recall no more words, read the list a second time with the following instructions: *"I am going to read the same list for a second time. Try to remember and tell me as many words as you can, including words you said the first time."* Put a check in the allocated space for each word the subject recalls after the second trial.
At the end of the second trial, inform the subject that (s)he will be asked to recall these words again by saying, *"I will ask you to recall those words again at the end of the test."*

Scoring: No points are given for Trials One and Two.

6. **Attention**:

Forward Digit Span: Administration: Give the following instruction: "*I am going to say some numbers and when I am through, repeat them to me exactly as I said them*". Read the five number sequence at a rate of one digit per second.

Backward Digit Span: Administration: Give the following instruction: "*Now I am going to say some more numbers, but when I am through you must repeat them to me in the backwards order.*" Read the three number sequence at a rate of one digit per second.

Scoring: Allocate one point for each sequence correctly repeated, (*N.B.*: the correct response for the backwards trial is 2-4-7).

Vigilance: Administration: The examiner reads the list of letters at a rate of one per second, after giving the following instruction: "*I am going to read a sequence of letters. Every time I say the letter A, tap your hand once. If I say a different letter, do not tap your hand*".

Scoring: Give one point if there is zero to one errors (an error is a tap on a wrong letter or a failure to tap on letter A).

Serial 7s: Administration: The examiner gives the following instruction: *"Now, I will ask you to count by subtracting seven from 100, and then, keep subtracting seven from your answer until I tell you to stop."* Give this instruction twice if necessary.

Scoring: This item is scored out of 3 points. Give no (0) points for no correct subtractions, 1 point for one correction subtraction, 2 points for two-to-three correct subtractions, and 3 points if the participant successfully makes four or five correct subtractions. Count each correct subtraction of 7 beginning at 100. Each subtraction is evaluated independently; that is, if the participant responds with an incorrect number but continues to correctly subtract 7 from it, give a point for each correct subtraction. For example, a participant may respond "92 – 85 – 78 – 71 – 64" where the "92" is incorrect, but all subsequent numbers are subtracted correctly. This is one error and the item would be given a score of 3.

7. Sentence repetition:

Administration: The examiner gives the following instructions: *"I am going to read you a sentence. Repeat it after me, exactly as I say it* [pause]: ***I only know that John is the one to help today."*** Following the response, say: *"Now I am going to read you another sentence. Repeat it after me, exactly as I say it* [pause]: ***The cat always hid under the couch when dogs were in the room."***

Scoring: Allocate 1 point for each sentence correctly repeated. Repetition must be exact. Be alert for errors that are omissions (e.g., omitting "only", "always") and substitutions/additions (e.g., "John is the one who helped today;" substituting "hides" for "hid", altering plurals, etc.).

8. Verbal fluency:

Administration: The examiner gives the following instruction: *"Tell me as many words as you can think of that begin with a certain letter of the alphabet that I will tell you in a moment. You can say any kind of word you want, except for proper nouns (like Bob or Boston), numbers, or words that begin with the same sound but have a different suffix, for example, love, lover, loving. I will tell you to stop after one minute. Are you ready?* [Pause] *Now, tell me as many words as you can think of that begin with the letter F.* [time for 60 sec]. *Stop."*

Scoring: Allocate one point if the subject generates 11 words or more in 60 sec. Record the subject's response in the bottom or side margins.

9. Abstraction:

Administration: The examiner asks the subject to explain what each pair of words has in common, starting with the example: *"Tell me how an orange and a banana are alike"*. If the subject answers in a concrete manner, then say only one additional time: *"Tell me another way in which those items are alike"*. If the subject does not give the appropriate response *(fruit)*, say, *"Yes, and they are also both fruit."* Do not give any additional instructions or clarification. After the practice trial, say: *"Now, tell me how a train and a bicycle are alike"*. Following the response, administer the second trial, saying: *"Now tell me how a ruler and a watch are alike"*. Do not give any additional instructions or prompts.

Scoring: Only the last two item pairs are scored. Give 1 point to each item pair correctly answered. The following responses are acceptable:

Train-bicycle = means of transportation, means of travelling, you take trips in both;

Ruler-watch = measuring instruments, used to measure.

The following responses are **not** acceptable: Train-bicycle = they have wheels; Ruler-watch = they have numbers.

10. Delayed recall:

Administration: The examiner gives the following instruction: *"I read some words to you earlier, which I asked you to remember. Tell me as many of those words as you can remember."* Make a check mark (√) for each of the words correctly recalled spontaneously without any cues, in the allocated space.

Scoring: **Allocate 1 point for each word recalled freely without any cues.**

Optional:

Following the delayed free recall trial, prompt the subject with the semantic category cue provided below for any word not recalled. Make a check mark (√) in the allocated space if the subject remembered the word with the help of a category or multiple-choice cue. Prompt all non-recalled words in this manner. If the subject does not recall the word after the category cue, give him/her a multiple choice trial, using the following example instruction, *"Which of the following words do you think it was, NOSE, FACE, or HAND?"*

Use the following category and/or multiple-choice cues for each word, when appropriate:

FACE:	category cue: part of the body	multiple choice: nose, face, hand
VELVET:	category cue: type of fabric	multiple choice: denim, cotton, velvet
CHURCH:	category cue: type of building	multiple choice: church, school, hospital
DAISY:	category cue: type of flower	multiple choice: rose, daisy, tulip
RED:	category cue: a colour	multiple choice: red, blue, green

Scoring: **No points are allocated for words recalled with a cue.** A cue is used for clinical information purposes only and can give the test interpreter additional information about the type of memory disorder. For memory deficits due to retrieval failures, performance can be improved with a cue. For memory deficits due to encoding failures, performance does not improve with a cue.

11. Orientation:

Administration: The examiner gives the following instructions: "Tell me the date today". If the subject does not give a complete answer, then prompt accordingly by saying: *"Tell me the [year, month, exact date, and day of the week]."* Then say: *"Now, tell me the name of this place, and which city it is in."*

Scoring: Give one point for each item correctly answered. The subject must tell the exact date and the exact place (name of hospital, clinic, office). No points are allocated if subject makes an error of one day for the day and date.

TOTAL SCORE: Sum all subscores listed on the right-hand side. Add one point for an individual who has 12 years or fewer of formal education, for a possible maximum of 30 points. A final total score of 26 and above is considered normal.

Appendix 1.1.6: Experiment 1 – Debrief

DEBRIEF FOR PARTICIPANTS.

Study: Word Processing and memory.

This study investigated the effects of word judging, as well as different amounts of reward in the form of money, on word learning. We tested this by creating different word judgment tasks as well as a rewarding, where you received money if you recognised the studied words correctly on the first memory test, and non-rewarding condition. As well as examining the effects of reward on intentional learning of word/images (when you were trying to memorise the materials because you knew your memory would be tested), we included a second surprise recognition test examining memory for the non-target, new words that were shown in the first recognition tests. This surprise test examined the effects of word judgment and reward on unintentional learning (when you learned new words/images during the first test without trying to do so). We were comparing the effects of word judgment and reward against strategic learning techniques to examine if there are differences in intentional and unintentional learning. We will look at the amount of words you correctly recognised from the two word judgment tasks and the high and low reward conditions, as well as how many new words you recognised in the final surprise test.

We anticipate that during the rewarding condition, participants will be able to recognise the studied words better on the first memory test. We also expect that participants will learn more new words during the reward memory test. We also anticipate that more strategic and elaborate encoding occurring during the semantic pleasantness judgment will lead to better recognition on the initial and surprise tests than less elaborate encoding during the vowel judgment task.

We hope that this study will help us understand better how reward can improve memory or memory strategy use, which could in the future lead to useful techniques for improving learning in educational settings or for people with memory problems.

Thank you very much for your important contribution to our research.

Please contact Louisa Salhi at the following email address; lmas6@kent.ac.uk or Dr Zara Bergstrom; Z.M.Bergstrom@kent.ac.uk if you have any questions regarding this study. Similarly if you wish to withdraw your information and contribution to the study please email with your personal created pin to withdraw the data. All of your information and data will be securely and anonymously stored.

Appendix 1.2: Chapter 2- Experiment 1 – Characteristics of Words

Table A.1.1. Length (no of letters), Syllable, Frequency, imageability and concreteness average ratings for the six word-lists. Frequency values represent Kucera-Frances written frequency (KFFRQ), imageability rating values represent (IMG) and Concreteness rating values (CNC).

Word List	Length	Syllables	KFFRQ	IMG	CNC
1	6.33 *(4-10)*	2 *(2-4)*	42.92 *(0-417)*	579.29 *(358- 631)*	590.12*(422- 653)*
2	6.13 *(3-10)*	1.96 *(1-4)*	33.33 *(1-348)*	578.66 *(521-624)*	588.2 *(545- 621)*
3	5.92 *(3-10)*	1.68 *(1-4)*	22.21 *(1-81)*	593.44 *(554-635)*	592.17 (516-639)
4	5.92 *(3-9)*	2.13 *(1-3)*	33.63 *(1-182)*	588.8 *(508-633)*	597.8 *(550-632)*
5	5.71 *(3-11)*	1.92 *(1-5)*	21.08 *(1-95)*	574.47 *(510-659)*	591.53 *(486-645)*
6	6.13 *(3-9)*	1.92 *(1-3)*	14.25 *(0-82)*	571.52 *(522- 628)*	591.57 *(522-644)*
7	5.25 *(3-8)*	1.75 *(1-3)*	27.92 *(3-98)*	582.05 *(486- 630)*	594 *(425-630)*
8	6.08 *(3-10)*	2.08 *(1-5)*	22.79 *(0-81)*	585.31 *(525- 639)*	581.06 *(487-615)*
9	5.88 *(4-10)*	1.83 *(1-4)*	57.75 *(0-591)*	586.2 *(481- 627)*	579.6 *(389-629)*
10	6.04 *(3-10)*	1.88 *(1-3)*	10.92 *(0-83)*	583.65 *(532- 619)*	590.24*(525-620)*
11	4.83 *(3-9)*	1.46 *(1-4)*	22.71 *(0-111)*	572.13 *(469- 635)*	588.04 *(521-637)*
12	6.08 *(4-9)*	1.88 *(0-3)*	18.79 *(0-99)*	586.38 *(533- 629)*	588.17 *(439-635)*

Scores are means with ranges for the average values (min-max) presented in parentheses.

Appendix 1.3: Chapter 2- Experiment 1 supplementary analysis

Appendix 1.3.1: Manipulation order effects on subsequent foil recognition

A novel design element in our study that was not present in the original paper by Jacoby et al. (2005) was the additional Reward manipulation, which we implemented as a within-participant independent variable across different blocks. Half the participants took part in the LOP manipulation before the Reward manipulation and the other half took part in the Reward manipulation before LOP manipulation, before their recognition of all foils was tested in a mixed surprise test. At the time of conducting each task, participants were always naïve to what other tasks would follow, but it is possible that their exposure to initial tasks changed how they conducted later tasks. The order of the LOP/Reward manipulation could therefore have affected the results, which could have contributed to differences between our findings and the original findings by Jacoby et al. (2005). We therefore conducted additional analyses to investigate if manipulation order had an effect on final foil recognition.

First, to investigate whether the LOP effect for foils was affected by order, we calculated the difference between Deep *Pr* and Shallow *Pr* on the Surprise test as shown in Table S3. This table shows that within the older group, the mean difference in *Pr* between Deep and Shallow foils was in fact numerically larger when the LOP phase was conducted *before* the Reward phase, suggesting that the intact source-constrained retrieval effect in our older group was not caused by the reward manipulation enhancing the use of source-constrained retrieval during the LOP phase (as this account would predict larger difference between deep and shallow foils when the LOP phase followed the Reward phase). Independent t-tests were conducted to compare the *Pr* Deep-Shallow difference measure between manipulation orders within each age group, which showed no significant effect of order within either group (Younger: $t(39)=1.00$, $p=.321$, $d= 0.31$; $BF_{10} =1/2.19$; Older: $t(39)=1.39$, $p=.172$, $d= 0.44$; $BF_{10} =1/1.52$), with relatively more support for the null hypothesis (no difference) than the alternative (a difference between order groups) from the Bayes Factors, although these Factors were not conclusive (likely due to the small group sizes available for this analysis).

Next, we tested whether the effect of Rewards on recognition of foils was modulated by order of the LOP and Reward manipulation blocks. Similarly, this analysis involved calculating the difference in *Pr* for Rewarded and Non-Rewarded foils on the Surprise test. The mean *Pr* difference values for this comparison (Table A.1.2) show that the reward manipulation had very little effect on foil recognition, regardless of order of the manipulation blocks. Independent t-tests were conducted within each age group, comparing the *Pr* difference measure across the manipulation order subgroups. There was no effect of order on the Reward-No Reward foil *Pr* difference within either group (Younger: *t*(39)=0.49, *p*=.629, *d*= 0.15; BF_{10} =1/2.97; Older: *t*(39)=0.08, *p*=.934, *d*=0.03; BF_{10} =1/3.26), with relatively more support for the null hypothesis (no difference) than the alternative (a difference between order groups) from the Bayes Factors.

Table A.1.2. *Mean and SD for the difference in Foil Pr split by foil manipulation type (Reward or LOP), age group and manipulation order (order of presentation of Reward and LOP blocks).*

Age group	Manipulation Order	Deep-Shallow *Pr* difference		Reward-No Reward *Pr* difference		
		M	*SD*	*M*	*SD*	*N*
Younger Adults	LOP-Reward	0.07	0.12	0.02	0.16	22
	Reward-LOP	0.12	0.16	<.001	0.15	19
Older Adults	LOP-Reward	0.13	0.16	-0.01	0.12	21
	Reward-LOP	0.07	0.14	<.001	0.16	20

Note: ‘manipulation order’ refers to whether participants conducted the LOP phase before the Reward phase, or vice versa.

Thus, importantly, there was no evidence from this analysis that participants were more or less likely to engage in source-constrained retrieval attempts or were more or less affected by Rewards when encoding foils depending on the order of LOP/Reward manipulation phases.

Appendix 1.3.2: Manipulation order effects on spontaneous strategy use during reward phases

Next, we examined if the order of LOP/Reward manipulations led to differences in spontaneous strategy use during the Reward study phases. For example, participants could have realised that deep processing was an effective learning strategy if completing the LOP tasks first, and this could have led them to also apply deep encoding strategies in the Reward study tasks.

Table A.1.3. *Percent of participants reporting spontaneous strategy use split by manipulation order, age group and type of Rewarded study task.*

Age Group	Manipulation Order	Reward	No Reward
Younger Adults	LOP-Reward	81	59
	Reward-LOP	84	79
Older Adults	LOP-Reward	67	52
	Reward-LOP	40	40

Note: 'Manipulation order' refers to whether participants conducted the LOP phase before the Reward phase, or vice versa.

Within each Age group, a chi^2 test was conducted to examine if self-reported strategy use (any strategy) during the Rewarded study phase differed depending on whether participants was exposed to LOP before Reward tasks, or vice versa. As seen in Table A.1.3, within the Older adults there was a non-significant tendency towards an increase in spontaneous strategy use during the Rewarded study phase for those who completed the LOP block first, compared to those who completed the Reward manipulation block first ($X(1)= 2.93$, $p=.087$, $phi=.27$) whereas within the Younger adults, strategy use in the Rewarded study phase did not differ dependent on manipulation order ($X(1)= 0.04$, $p=.839$, $phi=.032$). There were no significant differences in strategy use during the Non-Rewarded study phase depending on order of manipulation blocks for the Older adults ($X(1)= .631$, $p=.427$, $phi=.124$) or the Younger adults ($X(1)= 1.86$, $p=.173$, $phi=.213$).

Appendix 2: Chapter 3- Experiment 2

Appendix 2.1: Chapter 3- Experiment 2 – Experiment Paperwork

Appendix 2.1.1: Experiment 2 – Information Sheet

Study Information Sheet

Title of Project:	Word Judgments	**Ethics Approval Number:**	201614805240054138
Investigator(s):	Louisa Salhi Dr Zara Bergstrom	**Researcher Email:**	lmas6@kent.ac.uk

You are being invited to take part in a research study. Before you decide, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully. Please ask if anything is unclear, and do not feel rushed into making a decision.

Study Aim

This is a research study to examine how people remember information based on different amounts of reward, in the form of point values.

Eligibility Requirements

To participate in this study you must be over the age of 18 and have English as your native language. You must also have normal or corrected to normal vision (i.e. normal with glasses or contact lenses). Participants must not be Dyslexic. Participants need to be neurologically healthy and not currently taking any psychoactive medication such as anti-depressants.

What will happen to me in the study?

You will perform a task that involves memorising words that will appear on a computer screen. You will be given full instructions and get an opportunity to practice some elements of the task before the real experiment begins. After the study you may also be asked to complete some short questionnaires. The complete session will take around [30] minutes.

What will happen if I don't want to carry on with the study?

We want to emphasise that you may withdraw from the study at any time, and if you refuse to take part or decide to withdraw you will not suffer any penalty or loss of rights. If you do choose to withdraw or are no longer able to participate, then unless verbally directed otherwise the study investigators will keep the data collected up to that point. Your participation can be withdrawn by the study investigators.

Are there any negative side-effects/risks?

We do not anticipate any significant side-effects although some participants may feel a little tired after the test session.

What will happen to the results of the research study?

The results of this study will be used to better understand the mental processes that support word retrieval and how they are affected by rewards. The results of the study may be published for scientific purposes, but your records or identity will not be revealed unless required by law.

What happens to the information I provide?

Participation in this study guarantees confidentiality of the information you provide. No one apart from members of the research team and carefully selected research collaborators will have any access to the information you provide. Your name and any other identifying information will be stored separately from your data in a securely locked filing cabinet. Questionnaires will be stored in a securely locked room for as long as is required by the Data Protection Act, and then they will be destroyed by our confidential shredding service. You may withdraw from the study at any time, and if you refuse to take part or if you decide to withdraw, you will not suffer any penalty or loss of rights. Your participation in the study can be withdrawn by the study investigators.

What if there is a problem?

If you have a concern about any aspect of the study then you should speak with Dr Zara Bergström, who is director of the study. She can be reached on 01227 827507. If you remain unhappy and wish to complain formally, you can do this through the School of Psychology Chair of Ethics. Further details can be obtained from the School of Psychology General Office on 01227 824775.

Who is organising the research?

The research is organised by Louisa Salhi, a PhD student, under the supervisor of Dr Zara Bergström who works within the School of Psychology.

Who has reviewed this study?

The study has been approved by the School of Psychology, University of Kent Research Ethics committee.

If you would like a copy of the consent form or information sheet to keep, please ask the researcher. If you have any complaints or concerns about this research, you can direct these, in writing, to the Chair of the Psychology Research Ethics Committee by email at: psychethics@kent.ac.uk. Alternatively, you can contact us by post at: Ethics Committee Chair, School of Psychology, University of Kent, Canterbury, CT2 7NP.

Appendix 2.1.2: Experiment 2 – Consent Form

RESEARCH INFORMED CONSENT FORM

Title of Project:	Word Judgments	**Ethics Approval Number:**	201614805240054138
Investigator(s):	Louisa Salhi Dr Zara Bergström	**Researcher Email:**	lmas6@kent.ac.uk

Please read the following statements and, if you agree, initial the corresponding box to confirm agreement:
Signatures:

	Initials
I confirm that I have read and understand the information sheet for the above study. I have had the opportunity to consider the information, ask questions about the risks involved and have had these answered satisfactorily.	________
I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason.	________
I understand that my personal information and data will be treated as strictly confidential and will not be made public or shared with any other person outside of the research team.	________
I understand that my data may be published in a research article, but that no personal details will be divulged and that it will not be possible to identify any responses as my own.	
I confirm that I am over 16 years of age and that I freely agree to participate in this study.	

Name of participant (block capitals) ______________ Date ______________ Signature ______________

Researcher (block capitals) ______________ Date ______________ Signature ______________

If you would like a copy of this consent form to keep, please ask the researcher. If you have any complaints or concerns about this research, you can direct these, in writing, to the Chair of the Psychology Research Ethics Committee by email at: psychethics@kent.ac.uk. Alternatively, you can contact us by post at: Ethics Committee Chair, School of Psychology, University of Kent, Canterbury, CT2 7NP.

Appendix 2.1.3: Experiment 2 – Instructions

WELCOME TO THE WORD JUDGEMENT TASK

The experimenter will talk you through each stage.

Word judgment task 1:

Please can you judge the words in this list based on the vowles in the words.

- ➔ All we want you to do is identify if there is an "o" or "u" in the word.

- ➔ If there is press "Z key" to say yes
- ➔ If not press "M key" to say no.

- ➔ You will have around 2 seconds to make this judgments. So respond quickly but try to be as accurate as possible.

If you have any questions please ask.

PRESS SPACE BAR TO CONTINUE

[VOWEL JUDGEMENT TRIALS]

Visual Search Task 1:

You are now going to start the next task which is slightly different and getting you to judge images.

- ➔ During the next few trials you will see coloured pictures presented on the screen. We want you to look for a WHITE TRIANGLE that has been superimposed on top of some of the images.

We judge want you to look for a white triangle; " Is there a WHITE TRIANGLE on the screen?"

- ➔ YES: press the Z Key
- ➔ NO: press the M Key

Try to answer as quickly and accurately as possible.

If you have any questions please ask the experimenter, if not press 'space bar' to start.

[VISUAL SEARCH JUDGEMENT TRIALS]

Word Recognition 1:

Your next task is to do a word recognition task where you will receive POINTS for recognising words correctly!!! Your question will be “DO YOU RECOGNISE THE WORD?”

- ➔ You will see a cue before each word representing how many points you will win for **responding correctly to the upcoming word.** You will then have 2 seconds to respond to the word.

- ➔ The points will either be a **high value of 12 points** or a low value of 1 point
- ➔ Your aim is to win the maximum amount of points and beat your fellow participants!

- ➔ The top prize is a **£20.00 amazon voucher**, and there are two runner up prizes of £5.00 vouchers.

So for the next few trials you will see words presented on the screen. Some of these words are new and some you have seen before. DO YOU RECOGNISE THE WORD?

- ➔ YES it’s an OLD word by pressing the 'z' key
- ➔ NO it’s a NEW word by pressing the 'm' key

- ➔ **Remember you will receive points for CORRECTLY identifying a word as either "old" or "new".**
- ➔ Try your hardest to win the most points to win the maximum amount of POINTS!
- ➔ **The more accurate you are**, the more POINTS you will WIN !

- ➔ If you have any questions please ask. PRESS SPACE BAR TO START'

Try your hardest to WIN as many points as possible!

Please feel free to ask any questions. PRESS SPACE BAR TO CONTINUE to the practise trials,

- ➔ The instructions for these trials will be on the screen and this practise is here to give you an idea of how much time you have to respond to each word.

[PRACTISE TRIALS]

You are now ready to start the recognition task: *There will be a break mid-way through.*

- ➔ **Remember you will receive points for CORRECTLY identifying a word as either "old" or "new".**
- ➔ Try your hardest to win the most points to win the maximum amount of POINTS!
- ➔ **The more accurate you are**, the more POINTS you will WIN !

[RECOGNITION 1 TRIALS]

On screen

(Please tell the experimenter that you have reached this stage. They will instruct you about what to do next.)

Word Recognition 2

Now you are about to start the final test. For the next few trials you will see words presented on the screen.

- ➔ Some of these words are new and some you have seen before during the whole study.

DO YOU RECOGNISE THE WORD?

If you RECOGNISE any words from the WHOLE STUDY respond yes.

This includes all study and test phases so it can contains any words that have been presented on the screen so far!

- ➔ Use the Keyboard to respond by pressing:
- ➔ YES it’s an OLD word by pressing the 'z' key
- ➔ NO it’s a NEW word by pressing the 'm' key

If you have any questions please ask the experimenter.

When you are READY to start press SPACE BAR.

Appendix 2.1.4: Experiment 2 –Post Experiment Questionnaire

Strategy Questionnaire- *Word Memory.*

1: Did you use any techniques or strategies to help you **learn** the words in the first learning task? If so, please briefly tell us about these techniques. *(Please circle and write in the space below)*

YES NO

____

____.

2: Did you use any techniques or strategies to help you recognised the words in the **first test**? If so, please briefly tell us about these techniques and if they were different **between the high and low** value points. *(Please circle and write in the space below)*

YES NO Different technique

Comments:____

____.

3: How **motivated** were you to win the maximum amount of points during the **first test** *(Please circle and write in the space below)*

(not motivated) <- 1 2 3 4 5 -> (highly motivated)

Comments:____

____.

4: How **difficult** did you feel like the **first test** was to recognise the words you had earlier judged? *(Please circle and write in the space below)*

(not difficult) <- 1 2 3 4 5 -> (very difficult)

Comments:____

____.

5: How much **effort** did you put into recognising the words in the **first test** and winning the maximum amount of points during the **first test** *(Please circle and write in the space below)*

(no effort) <- 1 2 3 4 5 -> (high effort)

Comments:____

____.

Participant Id number:

6: Did you **anticipate** the final surprise test at the end? *(Please circle)*

YES NO

Comments:____
____.

7: Do you have any other comments about the study? *(Please write in the space below)*

Comments:____
____.

Appendix 2.1.5: Experiment 2 – Debrief

DEBRIEF FOR PARTICIPANTS.

Study: Word memory

This study investigated the effects of value on word retrieval and learning. We created high and low value conditions during the initial test phases, where different values were associated with different words. Your aim was to score as high as possible by recognising which words were old (previously studied) and which were new. You received points for correct identification of both old and new words.

The aim of this study was to examine if words associated with high value were recognised more accurately in the initial test compared to words associated with low value. As well as examining the effects of value on retrieval of words in the first test (when you were trying to maximise your score), we included a second surprise recognition test examining memory for the non-target, new words that were shown in the first recognition tests. This surprise test examined the effects of value on unintentional learning (when you learned new words during the first test without trying to do so). Similarly, we expected that during the surprise test, words associated with higher values when they were first seen in the previous test would be recognised more accurately than those associated to lower values.

We hope that this study will help us better understand how value can improve memory, which could in the future lead to useful techniques for improving learning in educational settings or for people with memory problems.

Thank you very much for your important contribution to our research.

Please contact Louisa Salhi at the following email address; lmas6@kent.ac.uk or Dr Zara Bergstrom; Z.M.Bergstrom@kent.ac.uk if you have any questions regarding this study. Similarly if you wish to withdraw your information and contribution to the study please email with your personal created pin to withdraw the data. All of your information and data will be securely and anonymously stored.

Appendix 3: Chapter 4- Experiment 3

Appendix 3.1: Chapter 4- Experiment 3 – Experiment Paperwork

Appendix 3.1.1: Experiment 3 – Information Sheet

Study Information Sheet

Title of Project:	Reward and memory.	**Ethics Approval Number:**	201614805240054138
Investigator(s):	Louisa Salhi Arianna Moccia Dr Zara Bergstrom	**Researcher Email:**	lmas6@kent.ac.uk am2210@kent.ac.uk

You are being invited to take part in a research study. Before you decide, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully. Please ask if anything is unclear, and do not feel rushed into making a decision.

Study Aim

This is a research study to examine how people learn and remember based on different amounts of reward, in the form of chocolate.

Eligibility Requirements

To participate in this study you must be over the age of 18 and have English as your native language. You must also have normal or corrected to normal vision (i.e. normal with glasses or contact lenses). Participants must not be Dyslexic. Participants need to be neurologically healthy and not currently taking any psychoactive medication such as anti-depressants. Participants must not be Diabetic and must also not be intolerant or allergic to milk chocolate.

What will happen to me in the study?

You will perform a task that involves memorising words and/or pictures that will appear on a computer screen. Before each task you will be told about what rewards you will receive for later remembering the words and/or images. Next, your memory for the words and/or pictures will be tested. After the study you will receive 10% of your winnings. The maximum amount you can win is 3 X 14g bag of chocolate buttons. You will be given full instructions and get an opportunity to practice some elements of the task before the real experiment begins. After the study you may also be asked to complete some short questionnaires. The complete session will take around 60 minutes.

What will happen if I don't want to carry on with the study?

We want to emphasise that you may withdraw from the study at any time, and if you refuse to take part or decide to withdraw you will not suffer any penalty or loss of rights. If you do choose to withdraw or are no longer able to participate, then unless verbally directed otherwise the study investigators will keep the data collected up to that point. Your participation can be withdrawn by the study investigators.

Are there any negative side-effects/risks?

We do not anticipate any significant side-effects although some participants may feel a little tired after the test session. There are some risks to participants if they are intolerant or allergic to milk chocolate without the participant's prior knowledge of this allergy; however you will not be required to consume any chocolate during the study itself.

What will happen to the results of the research study?

The results of this study will be used to better understand the mental processes that support word and/or picture memory and how they are affected by rewards. The results of the study may be published for scientific purposes, but your records or identity will not be revealed unless required by law.

What happens to the information I provide?

Participation in this study guarantees confidentiality of the information you provide. No one apart from members of the research team and carefully selected research collaborators will have any access to the information you provide. Your name and any other identifying information will be stored separately from your data in a securely locked filing cabinet. Questionnaires will be stored in a securely locked room for as long as is required by the Data Protection Act, and then they will be destroyed by our confidential shredding service. You may withdraw from the study at any time, and if you refuse to take part or if you decide to withdraw, you will not suffer any penalty or loss of rights. Your participation in the study can be withdrawn by the study investigators.

What if there is a problem?

If you have a concern about any aspect of the study then you should speak with Dr Zara Bergström, who is director of the study. She can be reached on 01227 827507. If you remain unhappy and wish to complain formally, you can do this through the School of Psychology Chair of Ethics. Further details can be obtained from the School of Psychology General Office on 01227 824775.

Who is organising the research?

The research is organised by Louisa Salhi, a PhD student, under the supervisor of Dr Zara Bergström who works within the School of Psychology.

Who has reviewed this study?

The study has been approved by the School of Psychology, University of Kent Research Ethics committee.

Appendix 3.1.2: Experiment 3 –Consent Form

RESEARCH INFORMED CONSENT FORM

Title of Project:	Reward and memory.	**Ethics Approval Number:**	201614805240054138
Investigator(s):	Louisa Salhi Arianna Moccia Dr Zara Bergström	**Researcher Email:**	lmas6@kent.ac.uk am2210@kent.ac.uk

Please read the following statements and, if you agree, initial the corresponding box to confirm agreement:

Signatures:

Initials

I confirm that I have read and understand the information sheet for the above study. I have had the opportunity to consider the information, ask questions about the risks involved and have had these answered satisfactorily.

I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason.

I understand that my personal information and data will be treated as strictly confidential and will not be made public or shared with any other person outside of the research team.

I understand that my data may be published in a research article, but that no personal details will be divulged and that it will not be possible to identify any responses as my own.

I confirm that I am over 18 years of age and that I freely agree to participate in this study.

Name of participant (block capitals) Date Signature

Researcher (block capitals) Date Signature

If you would like a copy of this consent form to keep, please ask the researcher. If you have any complaints or concerns about this research, you can direct these, in writing, to the Chair of the Psychology Research Ethics Committee by email at: psychethics@kent.ac.uk. Alternatively, you can contact us by post at: Ethics Committee Chair, School of Psychology, University of Kent, Canterbury, CT2 7NP.

Appendix 3.1.2: Experiment 3 Study Instructions

- **Memory 1_Instructions**

Please try to memorize the list of words; all the words will be used in the test.

For every correctly recognized word you will get 3 CHOCOLATE buttons!!!

If you have any questions please ask. If not press 'space bar' to continue.

- **Visual task 1_Instructions**

During the next few trials you will see pictures on the screen, and we want you to look for a white triangle. Not all the images will have a white triangle, but if you see one respond with the green key to say Yes there is a white triangle.

Is there a WHITE TRIANGLE on the screen?

Use the Keyboard to respond by pressing:

YES there is a WHITE TRIANGLE : press the green button

NO there is not a WHITE TRIANGLE : press the red button

Try to be as quick and accurate as possible.

If you have any questions please ask the experimenter, if not press 'space bar' to continue.

- **Memory 2_Instructions**

Please try to memorize the list of words; all the words will be used in the test.

For every correctly remembered word you will get 1/4 of a CHOCOLATE button.

If you have any questions please ask.

PRESS SPACE BAR TO CONTINUE

- **Visual task 2**

- **Memory Recognition _Instructions**

Now try to recognize which words were in the list and which words are new.

You will be then asked to use the **duration of the button press** to indicate your level of **confidence** in your response. We will finally tell you whether your response was correct or incorrect by showing you a **circle on the CORRECT response on the screen**.

Use the Keyboard to respond

'Z' Key -> 'old': a word from the list

'M' Key-> 'new': a new word

Hold down the keys to increase your confidence rating. The longer you hold the keys down, the more confident you are. This will show on the screen as a slow and **continuous color change** from **bright GREEN** to **bright RED.**

So if you are **very unsure** about the response you have given, you should press the key **very briefly and just once.** This might be not reflected in a colour change on the screen. The more you hold down the key the more the colour on the screen will **slowly and progressively change** from **GREEN** to **RED.**

For example if you are **QUITE confident** in your response, you should hold they key down until the colour changes to a graded pattern of **GREEN** and **RED**. If you are **DEFINITELY confident** in your response, you should then hold the key down until the colour changes to **bright RED**.

The computer will **ONLY record your first key press**, so please be sure to press the keys just once and hold them down to make the confidence rating without lifting your fingers from the buttons. **Please try to use the whole scale** and try not to just use the extremes of the scale or fall into any patterns. You will not have some **practice trials** where you can practice making unsure, quite confidence and highly confident judgments.

PRESS SPACE BAR TO CONTINUE TO THE PRACTICE TRIALS

- **Practice trials**

In the practice trials you will be asked to show the experimentor how you would make different old new confident judgments. You will be making these judgments for the same word so do not worry that the responses do not match up with the word.

If you are unsure about making these judgments or how to use the continuous scale please let the experimenter know before continuing.

1- Please show us how you would make a very confident ‘old’ judgment.
2- Now please show us how you would make a fairly confident ‘new’ judgment.

3- Now please show us how you would make a fairly unsure ‘new’ judgment.
4- Now please show us how you would make a very unsure ‘old’ judgment.
5- Finally please show us how you would make a quite confident ‘old’ judgment.

If you have any questions please ask the experimenter, otherwise press ‘space bar’ to continue.

- **Memory recognition 1_Instructions**

For every correctly recognized old and new word you will get 3 CHOCOLATE buttons!!!

Try to respond as accurately as possible using the Z (OLD) and M (NEW) keys and the continuous confident judgment scale.

Please feel free to ask any questions.

PRESS SPACE BAR TO CONTINUE

- **Visual task 3**

- **Memory recognition Instructions**

See memory recognition instructions above to refresh yourself of the old/ new confidence judgments.

- **Memory recognition 2_Instructions**

For every correctly recognized old or new word you will get 1/4 of a CHOCOLATE buttons.

Try to respond as accurately as possible using the Z (OLD) and M (NEW) keys and the continuous confident judgment scale.

Please feel free to ask any questions.

PRESS SPACE BAR TO CONTINUE

- **Visual task 4**

 # only provide once participants have reached the final test!!!

Surprise recognition _Instructions

In this part of the experiment you will see a word presented in the middle of the screen. You have seen some of these words before in the experiment already and some are completely new. You will be asked to make one of two kinds of recognition memory judgments: recollective or familiarity judgments or judge whether the word is new. **I will talk you through the difference between the recognition judgments first**.

Look at the instructions sheet and listen carefully while the experimenter explains them to you.

For example, sometimes you recognize someone's face, and perhaps remember talking to that person at a party the previous night. You may even recollect some details like what you were thinking while talking to that person, or which song was playing etc. This means that while recognizing this person you can actually RECOLLECT details about the event when you met her or him.

At other times, however, you recognize someone's face, you absolutely know that you have seen this person before, but you CANNOT recollect any contextual information as to where or when or how you have met this person before. So you have a strong feeling of familiarity in the absence of any recollection. This is the kind of distinction we are looking for.

Can you give me an example of a recollected memory?

Can you give me an example of a familiar memory?

PRESS THE SPACE BAR TO CONTINUE

Surprise_Instructions_2

So you have this 6-point scale with R, F1, F2, G, N1 and N2 and I will explain what each of them means.

R is the REMEMBER button. Please only press the REMEMBER button when you can recollect some qualitative information about the study event, that means some information from the context when you studied the word. For example, this could include such things as recollecting what you were thinking about when the word was presented, what the word looked or sounded like etc. If you remember my example this would be where you remember meeting that person at the party, maybe which song was playing etc. In general, you should only press a REMEMBER response if you could, if I asked you, tell me what you recollected about that study event.

If you cannot recollect anything specific about experiencing the item, you have other response options and I will talk you through these now.

F1 and F2 correspond to FAMILIARITY in the absence of any recollective experience - that is in the absence of contextual details. So, if you have a STRONG feeling of familiarity such as you have definitely seen this word before but you cannot recollect any details about the study event, then you should press F1. (Remember what I said before about the person you recognise and know that you know but you don't remember where you know the person from). If you think the word is QUITE familiar but you cannot recollect any details about the study event that is you don't remember any extra information about the study then you should press F2. You see that F1 and F2 are graded FAMILIARITY responses.

If you are not sure whether the word appeared or not at any point in the experiment and you will **guess the response, you should press G for GUESSING**.

And last, if you think the word was not presented before that is you have not seen it at any point during the experiment then you should press **N for NEW**. **N1 shows that you are UNSURE** that the response is NEW and **N2 that you are SURE** that the response is NEW.

R = REMEMBER

F1 = STRONG feelings of **FAMILIARITY**

F2 = QUITE FAMILIAR

G = GUESSING

N1 = UNSURE NEW

N2 = SURE NEW

(Say): PRESS THE SPACE BAR TO CONTINUE

Surprise_Instructions_3

This is how the scale looks like. You can see the buttons we are using on the keyboard, and they correspond to the R, F1, F2, G, N1, and N2 on the screen.

Please do try to use all 6 buttons (that is, don't fall into any patterns) and be as honest as you can.

You have enough time, in fact unlimited time, to think about your response and press the button.

1 = R

2 = F1

3 = F2

4 = F3

5 = N1

6 = N2

Just to make sure that the instructions are clear; could you please briefly explain which button means what?

PRESS THE SPACE BAR TO PROCEDE TO THE PRACTICE TRIALS

Appendix 3.1.3: Experiment 3 –Post Experiment Questionnaire

PARTICPANT ID:_____DATE:_/______/

Strategy Questionnaire

1: Did you anticipate the final surprise test at the end? *(Please circle)*

YES NO

2: Did you use any techniques or strategies to help you learn the words/pictures? If so, please briefly tell us about these techniques. *(Please circle and write in the space below)*

YES NO

____

____

____

____

3: Did you use different strategies to memorise or remember words/pictures during the high and low reward tasks? If so, please briefly tell us what you did differently. *(Please circle and write in the space below)*

YES NO

____

____

____

____

4: Do you have any other comments about the study? *(Please write in the space below)*

____

____

____

____

PARTICPANT ID:_____DATE:________/____/__.

Rate how much you like Milk Chocolate: *1- dislike 5- indifferent 10- love.*

1- Dislike, 2 3 4 5 6 7 8 9 10-love

Appendix 3.1.4: Experiment 3 – Debrief

DEBRIEF FOR PARTICIPANTS.

Study: Reward and memory.

This study investigated the effects of different reward levels in the form of chocolate, on word memory. We tested this by creating a high and a low reward condition, where you received different quantities of chocolate if you remembered the studied words correctly on the first memory test. As well as examining the effects of reward on intentional learning of word/images (when you were trying to memorise the materials because you knew your memory would be tested), we included a second surprise recognition test examining memory for the non-target, new words that were shown in the first recognition tests. This surprise test examined the effects of reward on unintentional learning (when you learned new words/images during the first test without trying to do so).

We will look at the amount of words/images you correctly recognised from the high and low value reward conditions, as well as how many new words you recognised in the final surprise test. We also asked you about how confident you were and about what type of memory experience you had for the words/ images in the final test (whether you felt you could vividly remember them or not). We will be looking at these judgements in relation to the reward value condition to see if rewards lead to different memory experiences.

We anticipate that during the high reward condition, participants will be able to recognise the studied words better on the first memory test. We also expect that participants will learn more new words during the high reward memory test and judge these recognition choices more confidently in the second surprise memory test.

We hope that this study will help us understand better how reward can improve memory, which could in the future lead to useful techniques for improving learning in educational settings or for people with memory problems.

Thank you very much for your important contribution to our research.

Please contact Louisa Salhi at the following email address; lmas6@kent.ac.uk or Dr Zara Bergstrom; Z.M.Bergstrom@kent.ac.uk if you have any questions regarding this study. Similarly if you wish to withdraw your information and contribution to the study please email with your personal created pin to withdraw the data. All of your information and data will be securely and anonymously stored.

Appendix 4: Chapter 5- EEG pilot and Experiment 4

Appendix 4.1: Chapter 5- EEG pilot study

Appendix 4.1.1: EEG pilot – EEG Information Sheet

EEG Study Information Sheet

Title of Project:	Electrical brain activity associated with word retrieval.	**Ethics Approval Number:**	201614805240054138
Investigator(s):	Louisa Salhi Dr Zara Bergstrom	**Researcher Email:**	lmas6@kent.ac.uk

You are being invited to take part in a research study. Before you decide, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully. Please ask if anything is unclear, and do not feel rushed into making a decision.

Study Aim

This is a research study to examine how people learn and remember words and how reward affects this, in the form of chocolate. We will also examine how the pattern of electrical activity within your brain varies dependent on how you process and remember words and pictures.

Eligibility Requirements- Can you take part?

To participate in this study, you must be between **the age of 18 and 35** and must have **English as your native language.** Participants must also have **normal or corrected to normal vision** (i.e. normal with glasses or contact lenses) and **be right-handed.**

Participants **must not have very thick hair**; must not have brides/ weave/ extensions. This will not allow us to fit the EEG cap effectively and we will not be able to record the EEG signal.

Participants must **not be Dyslexic**.

Participants need to be neurologically healthy and not currently taking any psychoactive medication such as anti-depressants.

Participants must also not be intolerant or allergic to milk chocolate; if so let the experimenter know the day before to organise dark chocolate substitute.

Before you come to the study?

- Try to get a good night sleep
- Eat breakfast or lunch and make sure you are not hungry
- Wash your hair the night before/ that morning
- Do not use/ leave in hair products such as gel/ oil/ mouse/ hairspray etc.
- Don't wear heavy make-up

What will happen to me in the study?

Before you begin the experiment, you will be fitted with a special cap that holds a large number of electrodes against your scalp. These electrode sensors will detect any changes in electrical brain activity (the electroencephalogram, or EEG) that is related to the task you will perform. A small number of loose electrodes will also be stuck onto your face and head with sticky tape to measure blinking and eye movements.

It is necessary to clean the area under each electrode site to ensure good electrode-skin contact. This may cause mild discomfort (due to the abrasiveness of the cleaner), and sometimes causes a little redness. To ensure good contact between the skin and all the electrodes, it is necessary to apply conductive, soluble gel which is best removed by washing your hair. The procedure should not be painful and is very safe. The application of electrodes usually takes around 30-45 minutes, but this may vary dependent on hair thickness.

Once the EEG cap has been attached, you will perform a task that involves memorising words that will appear on a computer screen. Before each task you will be told about what rewards you will receive for later remembering the words. Next, your memory for the words will be tested. After the study you will receive 10/20 % of your winnings. The maximum amount you can win is [2X 14 g bag of chocolate buttons]. You will be given full instructions and get an opportunity to practice some elements of the task before the real experiment begins. After the study you may also be asked to complete some short questionnaires. These tasks will take around 1hr 15 and the complete session will take around 2 ½ - 3 hours, including equipment set-up. You will have the opportunity to take breaks during the session.

<u>What will happen if I don't want to carry on with the study?</u>
We want to emphasise that you may withdraw from the study at any time, and if you refuse to take part or decide to withdraw you will not suffer any penalty or loss of rights. If you do choose to withdraw or are no longer able to participate, then unless verbally directed otherwise the study investigators will keep the data collected up to that point. Your participation can be withdrawn by the study investigators.

<u>Are there any negative side-effects/risks?</u>
We do not anticipate any significant side-effects although some participants may feel a little tired after the test session. There are some risks to participants if they are intolerant or allergic to milk chocolate without the participant's prior knowledge of this allergy; however you will not be required to consume any chocolate during the study itself.

<u>What will happen to the results of the research study?</u>
The results of this study will be used to better understand the brain processes that support word memory and how they are affected by rewards. The results of the study may be published for scientific purposes, but your records or identity will not be revealed unless required by law.

<u>What happens to the information I provide?</u>
Participation in this study guarantees confidentiality of the information you provide. No one apart from members of the research team and carefully selected research collaborators will have any access to the information you provide. Your name and any other identifying information will be stored separately from your data in a securely locked filing cabinet. Questionnaires will be stored in a securely locked room for as long as is required by the Data Protection Act, and then they will be destroyed by our confidential shredding service. You may withdraw from the study at any time, and if you refuse to take part or if you decide to withdraw, you will not suffer any penalty or loss of rights. Your participation in the study can be withdrawn by the study investigators.

<u>What if there is a problem?</u>
If you have a concern about any aspect of the study then you should speak with Dr Zara Bergström, who is director of the study. She can be reached on 01227 827507. If you remain unhappy and wish to complain formally, you can do this through the School of Psychology Chair of Ethics. Further details can be obtained from the School of Psychology General Office on 01227 824775.

<u>Who is organising the research?</u>
The research is organised by Louisa Salhi, a PhD student, under the supervisor of Dr Zara Bergström who works within the School of Psychology.

<u>Who has reviewed this study?</u> The study has been approved by the School of Psychology, University of Kent Research Ethics committee.

Appendix 4.1.2: EEG pilot write up

In this pilot study, I aimed to simplify the paradigm by removing the reward manipulation, taking a more parsimonious approach in the study design to investigate a clearer set of hypotheses. The intention of this study is to act as a pilot for the paradigm changes. These changes have been made to address firstly overcome the paradigm confounds related to response time caps which would have been interacting with response accuracy and confidence. Secondly, to provide test the appropriateness of using a new measure on the surprise test to examine source judgments instead Remember-Know responses. Tests of source memory provide more detailed information about the nature of memories, such as what contextual details people can remember about items in a recognition test (Bröder & Meiser, 2007; Johnson et al., 1993; Mitchell & Johnson, 2000)

The aim of this is to provide a better distinction between items encoded during the study test, and those that were possible falsely retrieved from the study phase. Previously the effect of false familiarity was speculative and the addition of source judgments will provide insight into these underlying recognition judgment processes. In sum, in this study I aimed to investigate if foils receiving false alarms on test 1 are later recognised more often than those receiving correct rejections on test 1, even after the removal of feedback. I also aimed to determine if this false alarm effect on foil memory was caused by a persistent false memory of studying those words, or if the increase in subsequent recognition was due to increased memory of encountering those items in the first test context. This issue was assessed by measuring source memory for both study and test phases on the final foil test. I also aimed to replicate the effect of fast reaction correct rejections on enhanced subsequent recognition, and address if a similar pattern occurs with an improved confidence analysis.

Based on Experiment 3's results, I expected to replicate the increase in subsequent recognition after a prior False Alarm compared to a prior Correct Rejection. I also expected that Fast-Correct Rejections would be recognised more often on a subsequent test compared to Slow-Correct Rejections. It was less clear how Confidence during Test 1 would affect subsequent recognition of foils, but I tentatively expected High Confident Correct Rejections to be recognised more often than Low Confident Correct Rejections, if novelty detection enhanced foil encoding. I used the Source

Judgments to test competing (although not mutually exclusive) hypotheses about why foils receiving false alarms on Test 1 were more often later recognised than foils that received correct rejections on Test 1. If it was the case that False Alarms were caused by persistent false memories across both tests, then I expected to find that prior False Alarms would be more often misattributed as having been seen during the Study phase than prior Correct Rejections. If false alarms however led to a boost in foil encoding (for example because of error monitoring/attentional processes), then source judgments for False Alarms should be more accurately attributed as having been seen during the Test phase.

Methods

Participants

Results are based on 17 native English-speaking students (4 Male; *M* age= 19.94 years, *SD*=4.94), two additional participants were excluded after low FA trial numbers (N<=10) and one other due to extremely slow response times. This study had a smaller sample size than Experiment 3 because of its confirmatory nature (aiming to primarily replicate the surprise test findings from Experiment 3), and because it was conducted to pilot an EEG compatible task design, that was to be subsequently used in the next experiment (presented in Chapter 5). The sample size enabled me to achieve ~0.7 power to detect an effect size of Cohen's *d*=0.55 (the effect size of the prior FA>prior CR difference in final foil recognition) at one-tailed α=.05.

The participants were recruited from the University of Kent Psychology participation scheme where students received course credit for their time. Participants had normal or corrected to normal vison and reported not having dyslexia, diabetes, or suffering from disorders like depression or anxiety and not taking any psychoactive medication. Importantly, participants had not taken part in any other experiments within this thesis. All participants provided full written consent and the procedures were approved by the University of Kent School of Psychology Ethics Board. Participants were tested by *LS.*

Materials

The experiment was programmed and run in PsychoPy (Peirce, 2007). Consistently with Experiment 3 the same word stimuli were used and again ordered in the same six lists (64 words in each). Word lists were counterbalanced across conditions, including the order of source judgments, which yielded a full counterbalance rotation of 12 participants.

The same milk chocolate food rewards as in Experiment 3 were provided, but this time as a constant reward to encourage task engagement (amount of chocolate was not manipulated). As in Experiment 3, 80 emotionally neutral images from the IAPS database (Lang et al., 2005) were used in a visual search filler task between encoding and recognition test phases.

Design and procedure

As with the previous experiments, all participants first completed an encoding task, here they were instructed to intentionally learn words, as with Experiment 3. Then participants completed an initial Old/New recognition task where new foils were incidentally encoded, and finally a surprise ‘Old/New’ recognition task for foils, which was adapted to include Source Judgments for ‘Old’ responses. All ‘Old/New’ recognition responses were given with continuous confidence measures as in Experiment 3. In brief to recap on paradigm alterations; changes to the task involved removing the reward manipulation so that a constant reward was delivered, removing feedback during the initial recognition test, removing the response time cap, and changing the format of the surprise test from a Remember-Know scale to an ‘Old/New’ judgment and follow up Source Judgments, both with continuous confidence measures. Like in previous experiments, in between each experimental stage a visual search filler task was used. This design created a fully within-participants factorial design. Counterbalancing of word list was controlled as well as source question presentation order.

a) Main recognition Task

The experimental procedure was modified from Experiment 3 procedures, which was based on the memory-for-foils paradigm, with an intentional study phase followed by an ‘Old/New’ recognition task and finally a surprise ‘Old/New’ recognition with additional Source Judgments as seen in Figure A.4.1.1.

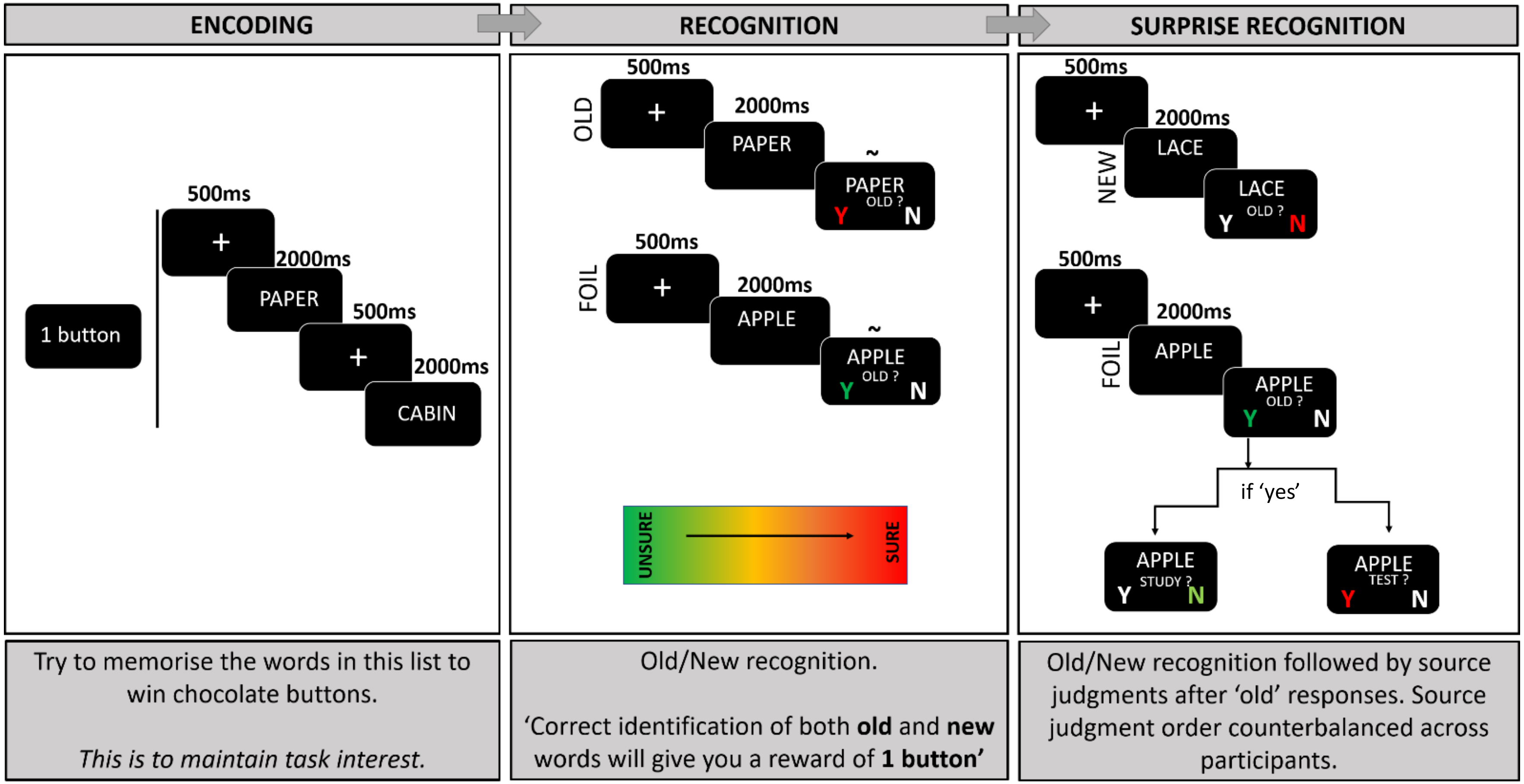


Figure A.4.1.1 Schematic of Pilot Study procedures. There were three phases, intentional encoding, recognition and a surprise recognition phase. Continuous confidence scores were made during the initial recognition tests and surprise tests. During the Surprise test source judgments were made for words that were responded to as 'Old'. Response button and 'yes'/ 'no' presentation order was counterbalanced across participants, as was source judgment order.

During the encoding phase participants were asked to intentionally try to learn a series of 128 words, shown one at a time for 2000ms after a 500ms fixation cross. Two word-lists made up the study phase (study block 1 and study block 2; 128 words in total), and words within each list were presented in a random order. Words were arranged in two lists to control the delay between study and test, so that the words from the first half of the study (study block 1) were tested first, and words from the second half of the study phase (study block 2) were always later tested after all the words from the first block. There was a break screen mid-way after 64 encoding trials to maintain the participant's attention and reduce fatigue. Participants were told that they would win 1 chocolate button every time they correctly later identified a word in the upcoming recognition phase. This constant reward was included in the study to try to keep the participants motivated and attentive throughout the lengthy study and test phases. Thus, the number of words to encode was identical to Experiment 3, but the trial timings were shorter, and no reward manipulation was used.

Next, participants proceeded to the Old/New recognition test phase where they were asked if they recognised words from the previous study phase or not. Previously seen 'Old' words were randomly intermixed with 'New' foil words, and participants were asked to respond to whether each word was 'Old?' by pressing buttons on the keyboard. The recognition test was made up of two test blocks (242 words in total) with 114 'Old' words and 128 'New' foil words. Importantly to note, 14 'Old' words that were studies were randomly selected to be presented in the Surprise Test only to legitimise the Source Judgment Questions (approx. 10% of words from both study lists, rounded up to the next integer). These 'Old' words were not presented in the test phase 1.

During the first test block, 'Old' words were always drawn from the first study block, and similarly during the second test block 'Old' words were always drawn from second study block, but this was only to maintain a standardised study-test delay and was not communicated to the participants. A break was given after every 64 trials, leading to 3 breaks within the recognition test phase.

During each recognition test trial, a fixation cross was presented for 500ms before a word was shown in the middle of the screen for 2000ms. Participants were told to not respond until the response options were shown after 2000ms, with the prompt question 'Old?' and the two options 'Yes' and 'No'.

Responses were made with the ‘z’ and the ‘m’ key. Allocation of these keys to ‘Yes’ and ‘No’ was fully counterbalanced. As in Experiment 3, participants were asked to indicate their confidence in their recognition decision by using the button press duration, with longer durations indicting higher confidence in their recognition decision. Participants received visual feedback in terms of continuous colour change from vivid green (Unsure) to vivid red (Very Sure) for their chosen option. When they released the button, the next trial started by presentation of the next fixation cross. Thus, in comparison to Experiment 2, trial timings were a bit shorter, and the response window was self-paced. This important change was implemented to removed potential response deadline confounds discussed in with regards to Experiment 3, meaning that participants could now always use the full confidence scale as there was no upper time limit to respond. There was an extensive verbal practise session provided before the test to familiarise participants with the confidence scale and key responses.

Then participants were told about the surprise recognition test, where they were instructed that words would be shown on the screen and these could have either appeared at any time during the experiment so far, or could be completely new. Overall, during the surprise recognition test participants saw 270 words in total. They were presented with all previously seen foil words from the recognition test 1 phase (128 foil words, 64 from test block 1 and 64 from block 2) as well as 10% of the studied items from the study phase (14 studied words, 7 from study block 1 and 2) and new words (128 unseen words, two new lists of 64 words each). The studied words (14) were included to legitimise the Source Judgment options. Therefore, by including ‘studied’ items that had not been presented in the first test phase there was a genuine response option whereby you would be able to respond ‘Yes’ to seeing the word in the ‘Study’ phase, but ‘No’ to seeing the word in the ‘Test’ phase. The aim of this was to prevent participants always responding ‘Yes’ to both the ‘Study’ and ‘Test’ options whenever they recognised the word only. If all the studied words had been included in the first test phase, then this would not have been a genuine response option, and could have biased the participants response tendencies. Importantly, participants were unaware of the probabilities of word conditions within this phase and were told any words could appear.

Participants were tasked to make a recognition judgment whether they had seen the word before at any point during the prior tasks. Participants used the ‘z’ and ‘m’ keys to respond ‘Yes’ or ‘No’ to the question; ‘Is the word ‘Old?’. Key assignment to responses were counterbalanced but kept consistent across tasks within participants, and were indicated by yes/no options on the screen. Similar to in recognition Test 1, participants also provided their confidence alongside their recognition responses by using their button press duration, indicated on the screen with a colour change (as described previously). Once again participants had unlimited time to respond during the surprise test. This judgment was thus changed from the Remember/Know/New judgements in Experiment 3.

Another addition as compared to Experiment 3 was that if participants responded that the word was “Old”, two follow up source questions were asked; ‘Did you see this word in the study phase?’ and ‘Did you see this word in the test phase?’ (The order of these questions was counterbalanced across different participants, but kept constant within participants). Again, they selected ‘Yes’ or ‘No’ using the same two keys as for the ‘Old?’ judgement, and indicated their confidence in the chosen response using the button press durations. Extensive, verbal and written instructions were given to explain the source judgments and practise was given in the presence of the experimenter to ensure participants understood the difference between each phase and how to respond to the source judgment questions. For example, they practiced how to respond when they had only studied but not been tested on a word, compared to when they had studied and been tested on that word. Multiple scenarios and responses were practised with the experimenter talking through each practise trial, with the aim to show participants that they should respond to the study and test source questions independently.

b) Visual Search filler task

In between all experimental phases a visual search task was used as a filler to prevent rehearsal between each experimental task (two times in total after the encoding phase and the again after the recognition test phase). The task was exactly the same as in prior chapters and lasted for 1 minute in total. See Chapter 2 for detailed description of the Visual search task used.

c) Post-Experiment Debrief

All participants were given verbal and written debrief and were given 2 bags of milk chocolate buttons or lactose-free equivalent (45g). All participants were given the same amount of chocolate winnings.

Data Analysis

For the Test 1 phase, the proportion of correct responses were calculated separately for Old words (i.e. hit rates) and New words (i.e. correct rejection rates). Discrimination indices (*Pr*) and Response bias (*Br*) have been calculated using the above raw scores. This provides a comparable level of accuracy as compared with my prior experiments. *Pr* and *Br* were calculated for Test 1 in an identical way to Chapter 2 (see 2.3.4 for details). Mean reaction times (RTs) and continuous Confidence judgments were also calculated for Old and New words, split by response accuracy (Hit/Miss and CR/FA). The analysis conducted on the recognition Test 1 validated the different trial conditions that were used to assess the rate of incidental encoding occurring during test 1, as measured on the subsequent surprise recognition test.

On the Surprise Test, in addition to recognition accuracy and Source Judgment accuracy, surprise test Confidence measures were also assessed. Reaction Times were not analysed on this test since participants were told to focus on accurate rather than fast decision making. On the Surprise Test, an overall measure of Discrimination was calculated as well as an overall Bias measure. Discrimination and Bias related to Test 1 Response Times splits or Confidence splits were not calculated due to the conditions being conditionalized on participant responses. These scores would therefore iterate on a within participant level and would not be any more meaningful that the raw scores given that 'New' items on the surprise test would hold the same proportional value.

So therefore, similarly to in the latter analysis for Experiment 3, the proportion of accurate responses were calculated separately for each foil conditions based on prior test 1 Foil accuracy (Prior False Alarm (pFA); Prior Correct Reject (pCR)), and for New items (resulting in hit rates for foils since these items are "old" on the final test, and correct Rejection rates for New items). Experiment

3's findings stimulated this experiment and therefore I had directional hypothesis about how Prior Correct Rejections would be responded to in the Surprise Test in regard to Response times and Confidence on Test 1. The Correct Rejection (CR) trial splitting was conducted in the same way as in Experiment 3, with a median value for each person being calculated separately for Reaction Times and Confidence. Then participant's CRs were categorised into either below the median value or above the median value ensuring an equal split of Correct Rejection trials. This yields four additional trial conditions, those based on Reaction Times; Fast and Slow Correct Rejections, and those based on Confidence; High and Low Confident Correct Rejections. These will then be used in comparisons to prior False Alarm trials.

Source judgments will be calculated based on the proportion of correct responses to the two questions; "Study?" and "Test?". These responses were categorised based on prior Test 1 responses, Prior FA, Prior CR. Confidence scores were also calculated for these Source Responses. A similar analysis will be conducted on the Correct Rejection splits (by Reaction times and by Confidence) in the full EEG study.

In this study the main aims were to examine if the paradigm changes yielding replicable effects to those in Experiment 3, and that source judgments were appropriate to be used. This will be examined by looking at source judgments for FA and CR's only in this study.

For details of the frequentist inferential statistical tests and the Bayesian statistics tests refer to detailed descriptions in Chapter 2 (see 2.3.4).

Results

Recognition Test 1

As seen in Table A.4.1.1, accuracy at the first Old/New recognition test in this pilot study was similar to in Experiment 3, for both Old and New Foil items. Overall Test 1 Discrimination ability was good (*Pr; M*= .41, *SD*=.15) and Response Bias was neutral (*Br; M*= .52, *SD*=.12).

There was also a clear difference in Response Times and Confidence between correct and incorrect responses, which is in line with Experiment 3. This pattern indicates that the confidence scale was used effectively, and together with the removal of the 2700ms response cap enabled me to assess the effect of Test 1 foil judgement confidence and response times on subsequent foil recognition without confounding these two variables.

Table A.4.1.1: *Descriptive Statistics of Test 1 Old and New words split by Test 1 Accuracy.*

Response Type	Proportion		Response Time		Confidence	
Hit	71	(11)	533	(143)	1473	(449)
Miss	-	-	795	(269)	788	(468)
Correct Rejection	70	(09)	692	(176)	942	(444)
False Alarm	-	-	825	(381)	748	(345)

Note. Mean and SD shown in parenthesis, Response time and Confidence are shown in *ms*. N= 17

It was also important to check that the median split calculations were successful to yield Correct Rejection splits by Response Time (Fast and Slow) and of Confidence (High and Low), and to assess how these conditions related to Response Times and Confidence of the False Alarms, as these trial splits were to create variables for the Surprise Test analysis. As a recap; I created these conditions by calculating a median value for Confidence for each participant's Correct Rejection trials. This value was then used to classify that participant's Correct Rejection trials into either High Confident (above their median) or Low Confident (below their median) and a similar median split was conducted separately based on Response Times.

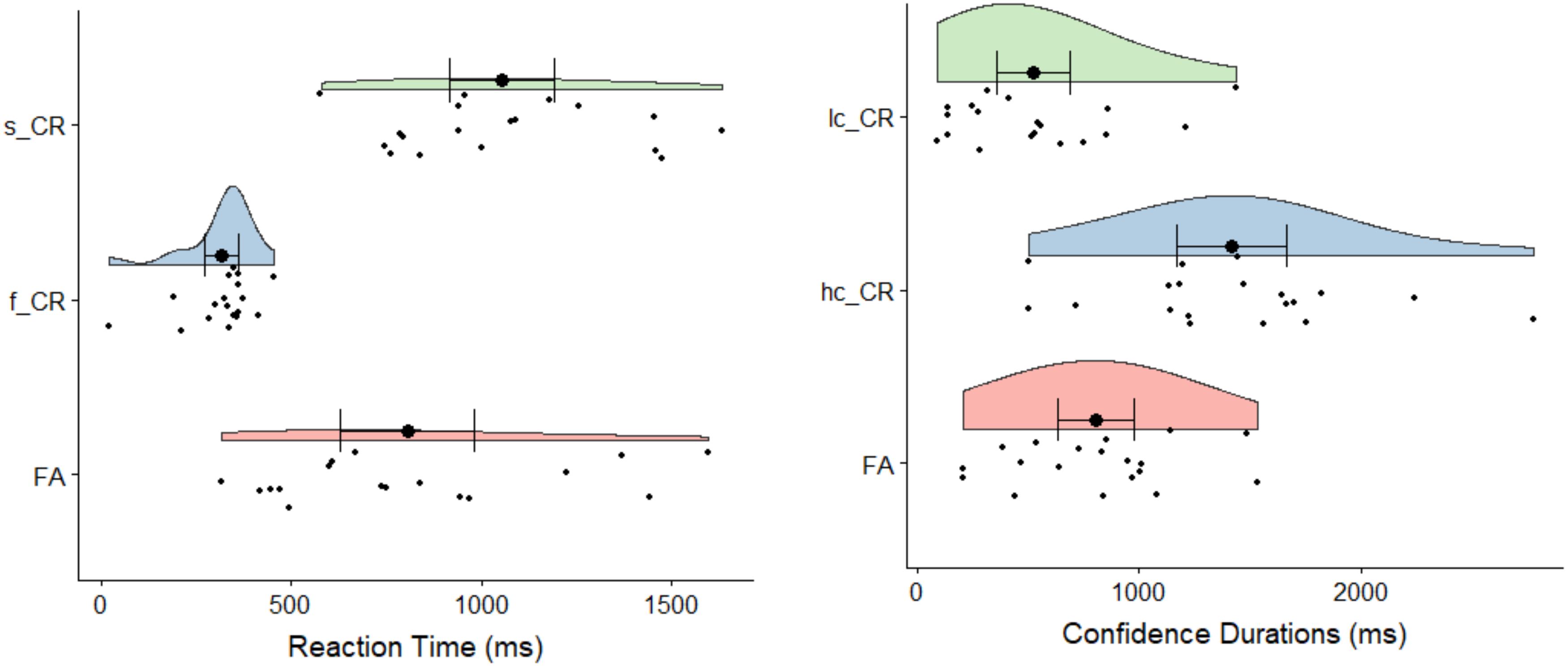


Figure A.4.1.2. Distributions of Response times and Confidence during Recognition Test 1. The plot on the left shows Response Times (sec) and the plot on the right shows Confidence durations (sec). The plot is for Foil items, split by Accuracy (FA and CR) and for Correct Rejections further split again by either; Fast and Slow Response Times (right plot) or by High and Low Confident decisions (left plot

Table A.4.1.2: *Descriptive Statistics of Test 1 Response times and Confidence for New Foil words split by Test 1 Accuracy and separately for Correct Rejection response time and confidence.*

Measure	Word Type	Time Durations (*ms*)
Response Time	Slow half CR	1071 (298)
	Fast half CR	316 (98)
	FA	825 (381)
Confidence	Low half CR	495 (350)
	High half CR	1389 (580)
	FA	748 (345)

Note. Mean and SD shown in parenthesis, Response time and Confidence shown in *ms*. N= 17

As seen in Figure A.4.1.2 and Table A.4.1.2, the median split calculations have yielded a clear difference between both High (hc.CR) and Low Confident Correct Rejection (lc.CR) trials and Slow (s.CR) and Fast Correct Rejection (f.CR) trials. False Alarm values for Response Times and Confidence durations importantly remained between these Correct Rejection split values. There was a significant main effect of Foil Type on test 1 response times ($F(2,32)=63.78$, $p<.001$, $n_p^2=.799$), this was similarly the case for test 1 confidence ($F(2,32)=68.08$, $p<.001$, $n_p^2=.810$). All follow up pairwise t-tests indicated significant differences between test 1 Response Times and Confidence as seen in Table A.4.1.3.

Table A.4.1.3. *Paired-samples t-tests between False Alarms and split conditions for Correct Rejects during Recognition Test 1 based on Response Time (Fast and Slow Responses) and separately based on confidence (High and Low Confidence).*

	Paired tests	*t*	*p*	*d*	BF_{10}
Response Times	FA x sCR	**4.58**	**<.001**	**0.72**	**105.70**
	FA x fCR	**6.18**	**<.001**	**1.83**	**1712.46**
	fCR x sCR	**11.55**	**<.001**	**3.40**	**3.116** $^{\text{e+6}}$
Confidence	FA x lcCR	**4.67**	**<.001**	**0.73**	**123.77**
	FA x hcCR	**7.14**	**<.001**	**1.34**	**8178.51**
	hcCR x lcCR	**10.18**	**<.001**	**1.87**	**594533.87**

Note. Significant values below .05 are shown in bold. N=17.

Surprise test results

Overall accuracy on the surprise test was high (*M*=.82, *SD*=.15). Discrimination ability (Pr) was good (*Pr; M*=.54, *SD*=.17) and response bias was on the conservative side of neutral (*Br; M*=.64, *SD*=.20)

Subsequent recognition given Test 1 Accuracy

Overall when comparing subsequent recognition based on Test 1 accuracy, there was a clear enhancement of subsequent recognition for prior False Alarms compared to prior Correct Rejections (*t*(16)=4.17, *p*=.001, *d*=0.43, BF_{10}=50.47) seen in Table A.4.1.4 and replicating findings in Experiment 3.

In contrast with Experiment 3, this difference in subsequent recognition was not accompanied by differences in Confidence, with no differences seen between Confidence ratings for prior False Alarms compared to prior Correct Rejections (*t*(16)=1.33, *p*=.203, *d*=0.11, BF_{10}=1/1.90).

Table A.4.1.4. *Proportion accurate responses on the Surprise Recognition test for new items and foil items. Foil item accuracy is presented split by Test 1 response accuracy, and for prior correct rejections split again separately by fast and slow response times and then by high and low confidence on Test 1.*

Word Type	Surprise Test Proportion Accuracy
New	72 (12)
Prior FA	87 (14)
Prior CR	80 (17)
Prior s.CR	81 (18)
Prior f.CR	80 (16)
Prior lc.CR	78 (17)
Prior hc.CR	82 (18)

Note. Mean and SD shown in parenthesis. F.CR represents fast CR half, s.CR represents slow CR half; hc.CR represents high confident Cr half, lc.CR represents low confident CR half. *N*=17

Subsequent recognition Source judgments

Next, I analysed Source Judgments for Foil items that received and "Old" recognition response, depending on if the response to the Foil was previously incorrect or correct on Test 1, to test the processes driving subsequent recognition enhancements of False Alarms compared to Correct Rejections. If enhanced recognition of prior FAs was due to persistent false memory of studying those foils, I expected to see an increase in inaccurate endorsement of having seen those items in the Study phase (i.e. inaccurate responses to the 'Study?' source question). In contrast, if prior FAs were associated with enhanced encoding (for example due to increased monitoring of items that led to incorrect responses) I would expect to see an increase in accurate endorsement of having seen those items in the Test phase (i.e. accurate responses to the 'Test?' source question).

Table A.4.1.5. *Source Judgment responses to foil items that were previously a False Alarm or Correct Rejection and subsequently correctly recognised on the Surprise Test.*

Word Type	Proportion Responses	
	Study?	Test?
Prior FA	61 (25)	84 (22)
Prior CR	34 (34)	81 (29)

Note. Mean and SD in parenthesis shown for source question yes responses. N= 17

As shown in Table A.4.1.5; there was a clear increase in the proportion of incorrect responses to the 'Study?' question for prior False Alarms compared to prior Correct Rejections (*t*(16)=4.54, *p*<.001, *d*=.90, BF_{10}=97.64). In contrast, there was no difference between prior False Alarms and Correct Rejections in correct responses to the 'Test?' question (*t*(16)=0.949, *p*=.327, *d*=.01, BF_{10} =1/2.71).

As participants also gave Confidence ratings for Source Judgments this provides richer detail about the Source Judgment decisions, as shown in Table A.4.1.6.

Table A.4.1.6. *Confidence Ratings (in ms) for the "Study" and "Test" Source Questions. Confidence has been split by Response type (yes/no) and Prior Foil Type (prior FA or prior CR).*

Confidence Ratings (*ms*)			
Source Question	Response	FA	CR
Study? N=14	Yes	900 (318)	658 (321)
	No	684 (325)	888 (309)
Test? N=11	Yes	1115 (247)	1169 (300)
	No	841 (481)	757 (536)

Note- Mean and SD in parenthesis. N for descriptives presented in line with pairwise exclusion used in ANOVA, hence variations in N.

As shown in Table A.4.1.7, there was an interaction between Foil Type (prior FA vs. prior CR) and Source Question Response (Yes vs. No) for Confidence ratings during the ‘Study?’ Source Judgments, but there was no such interaction for ‘Test?’ Source Judgements (there was only a simple main effect of Response, with higher confidence for ‘Yes’ than ‘No’ responses).

Table A.4.1.7. *Results of 2x2ANOVAs examining differences in Source judgement Confidence ratings as a function of Source Responses and prior Foil accuracy.*

2 x 2 ANOVA: Foil Type (prior FA x prior CR) X Source Qu Response (Yes x N0)				
Source Question		*F*	*p*	*np2*
Confidence to "Study" Question Response	Foil Type	0.20	.663	.015
	Source Response	<0.01	.924	.001
	Foil Type*Source Response	**43.96**	**<.001**	**.772**
	N=14			
Confidence to "Test" Question Response	Foil Type	0.06	.815	.006
	Source Response	**5.09**	**.048**	**.337**
	Foil Type*Source Response	0.89	.367	.082
	N=11			

Note. Significant values below .05 are shown in bold.

Follow up tests showed that for the ‘Study’ source judgement, Confidence in the response was higher for prior False Alarms when participants responded ‘Yes’ compared to ‘No’ (*t*(13)=2.37, *p*=.034, *d*= 0.67, BF$_{10}$= 2.14), whereas Confidence was higher for prior Correct Rejections when responding ‘No’ compared to ‘Yes’ (*t*(13)=3.87, *p*=.002, *d*= 0.73, BF$_{10}$=22.47). Confidence in the Study question was higher for prior False Alarms compared to prior Correct Rejections when responding ‘Yes’ (*t*(16)=2.43, *p*=.027, *d*= 0.36, BF$_{10}$=2.38). In contrast, Confidence was significantly lower when responding ‘No’ to the Study question for prior False Alarms compared to prior Correct Rejections (*t*(13)=5.07, *p*<.001, *d*= 0.64, BF$_{10}$=148.37) even though ‘No’ was the correct response.

Thus, both source judgements and associated confidence ratings showed that participants were more likely to falsely believe a foil had been shown in the study phase if that foil had also received a false alarm in the first test, compared to foils that received correct rejections on the first test.

Subsequent recognition given Test 1 Accuracy and Test 1 CR Response Time

In the next analysis, I aimed to replicate the finding from Experiment 3 that foils that received fast correct rejections on Test 1 were, similarly to foils receiving false alarms, also more likely to be recognised on the surprise test, compared to foils receiving slow prior Correct Rejections (see Figure A.4.3 for descriptive statistics). Similar to in Experiment 3, the raw hits will be used for the following analysis. This is because the *Pr* and *Br* scores would be redundant as the proportion is calculated using the same new response rates, having only one set of 'New' items. Therefore, for consistency with Experiment 3 and transparency, raw scores will be used for the analysis.

There was a strong increase in subsequent recognition in the surprise test for prior False Alarms compared to prior Fast Correct Rejections (f.CR) (*t*(16)=3.91, *p*<.001, *d*= 0.45, BF_{10}=31.89) and a weaker, yet still significant increase in subsequent recognition for prior False Alarms compared to prior Slow Correct Rejections (s.CR) (*t*(16)=2.86, *p*=.011, *d*= 0.36, BF_{10}=4.86). This result is thus inconsistent with the findings in Experiment 3. No significant differences were seen between f.CR and s.CR (*t*(16)=0.47, *p*=.648, *d*= 0.06, BF_{10}=1/3.65).

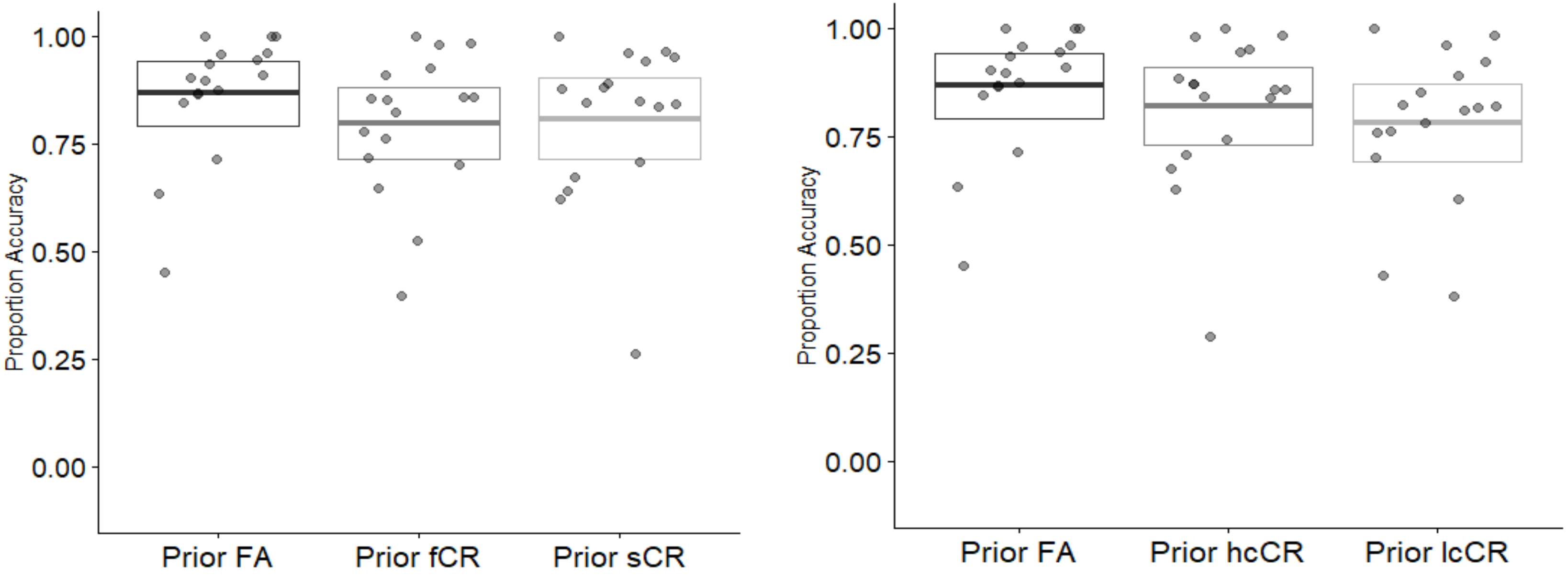


Figure A.4.3: Proportion accuracy distributions during the Surprise Test for Foils, split by Prior Accuracy in Test 1. The plot on the left shows surprise test proportion accuracy results for words that were a prior False Alarms and prior Correct Rejections which were split by Fast (f.CR) and Slow (s.CR) Response Times in Test 1. The plot on the right shows surprise test proportion accuracy results for words that were a prior False Alarms and prior Correct Rejections which were split by High (hc.CR) and Low (lc.CR) Confidence durations in Test 1. The dot plot shows each individual score, with the mean value and confidence interval plotted as a centre line and box respectively. Scores have been randomly scattered across the x-axis for visualisation purposes. *N*=1

Subsequent recognition given Test 1 Accuracy and Test 1 CR Confidence

Subsequent recognition in the surprise test was significantly higher for prior False Alarms compared to both High Confident Correct Rejections (hc.CR) (*t*(16)=2.40, *p*=.029, *d*= 0.30, BF$_{10}$=2.26) and Low Confident Correct Rejections (lc.CR) (*t*(16)=4.28, *p*=.001, *d*= 0.53, BF$_{10}$=61.36). This difference was not apparent in Experiment 3 where no clear differences were seen in the Correct Rejections made with either High or Low Confidence compared to False Alarms. There was however, no significant differences in subsequent recognition between hc.CR and lc.CR (*t*(16)=1.50, *p*=.153, *d*=0.21, BF$_{10}$=1/1.56).

Discussion

This experiment was conducted to address unanswered questions that arose in Experiment 3 and to pilot paradigm changes before running a large EEG study (Experiment 4 in Chapter 5). Unexpectedly in Experiment 3, there was an increase in subsequent recognition after a false alarm error and after rapidly identifying a new item, compared to more slowly on an initial recognition task. These findings raised questions about what cognitive processes related to accuracy and processing speed influenced recognition retrieval attempts and incidental foil encoding. However, as these effects were not hypothesised, the paradigm was not optimised to investigate these effects and it was also not clear if the effects were replicable. Therefore, in this pilot study I aimed to more directly investigate the processes during Test 1 that related to accuracy and error monitoring, and which resulted in increased subsequent recognition of foils. I therefore made a number of paradigm changes to more directly test a priori predictions. I removed the redundant reward manipulation, changed response requirements to allow for self-paced responses to better disentangle reaction times and confidence, and added source judgments to the surprise test to assess the underlying cognitive processes that resulted in enhanced recognition of prior false alarms on the final test.

In this pilot study, the increased subsequent recognition for foils receiving false alarms over correct rejections on Test 1 was replicated. Source judgements provided a new insight into reasons driving this strong effect, which was seen even in a much more limited smaller sample. The source judgements suggest that enhanced surprise recognition of foils receiving prior false alarm responses were driven by persistent false memories of studying those items, as there was a clear increase in prior false alarm foils being incorrectly attributed to the study phase, even though those items were only encountered in the test phase. However, the prior false alarm and prior correct rejection conditions did not differ in source attributions to the test phase, therefore there was no evidence that these items differed in encoding of the test context where they were actually encountered. This result is interesting, and may be due to spontaneous false memory creation, similar to that observed in the Deese-Roediger-McDermott (DRM) paradigm (Roediger & McDermott, 1995). In the DRM paradigm studied items such as word lists and 'extra-experimental' familiarity of the stimuli in question influence the creation of false memories (Jou & Flores, 2013; Mcdermott & Roediger, 1998). In the DRM paradigm lists of words are used to prime false memory activation through activating semantic associates. For example, presenting the words “bed, rest, tired, dream, nap” would activate the semantically associated word “sleep”. In this example, it would be incorrect to say you have seen the word ‘sleep’ before, but commonly having seen the list of associated words, this will lead to a false memory of seeing the word ‘sleep’. Linking this back to the current findings, there could be semantically associated words within the study lists. This could help explain why some words illicit a consistent false memory (False Alarms) and a source memory of studying those words, even if they have never encountered them in the task before.

Unlike in the DRM paradigm where sets of words are primed intentionally, in this experiment, the words were not combined in a certain way to illicit false memories. Therefore, each individual may semantically associate words differently, and have differing levels of ‘extra-experimental’ false-familiarity with those items. For example, if they recently read a newspaper article before the study with key words closely associated, this could illicit a false memory of studying those items or associated items in the actual study. Importantly, this would differ from person to person, and is interesting as it provides a naturalistic view on what processes are involved in recognition abilities and how spontaneous false memories affect recognition success.

This pilot study was primarily conducted to test the adapted paradigm for future use with EEG (see Chapter 5) and therefore had limited statistical power with a small sample size (N=17). The effects of Response Time and Confidence on subsequent recognition were not consistent across both studies and this will need to be assessed further before making generalisable conclusions. Importantly however, with the paradigm alterations there was a clear replication of the effect of false alarms compared to correct rejections, which was the key effect of interest and provides a suitable paradigm to apply in my next experiment using EEG.

Appendix 4.2: Chapter 5- EEG study

Appendix 4.2.1: EEG study paperwork

Appendix 4.2.1.1: Experiment 4- Information sheet

EEG Study Information Sheet

Title of Project:	Electrical brain activity associated with word retrieval.	**Ethics Approval Number:**	201614805240054138
Investigator(s):	Louisa Salhi Dr Zara Bergstrom	**Researcher Email:**	lmas6@kent.ac.uk

You are being invited to take part in a research study. Before you decide, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully. Please ask if anything is unclear, and do not feel rushed into making a decision.

Study Aim

This is a research study to examine how people learn and remember words and how reward affects this, in the form of chocolate. We will also examine how the pattern of electrical activity within your brain varies dependent on how you process and remember words and pictures.

Eligibility Requirements- Can you take part?

To participate in this study, you must be **over the age of 60** and must have **English as your native language.** Participants must also have **normal or corrected to normal vision** (i.e. normal with glasses or contact lenses) and **be right-handed.**

Participants **must not have very thick hair**; must not have brides/ weave/ extensions. This will not allow us to fit the EEG cap effectively and we will not be able to record the EEG signal.

Participants must **not be Dyslexic**.

Participants need to be neurologically healthy and not currently taking any psychoactive medication such as anti-depressants.

Participants must also not be intolerant or allergic to milk chocolate; if so let the experimenter know the day before to organise dark chocolate substitute.

Before you come to the study?

- Try to get a good night sleep
- Eat breakfast or lunch and make sure you are not hungry
- Wash your hair the night before/ that morning
- Do not use/ leave in hair products such as gel/ oil/ mouse/ hairspray etc.
- Don't wear heavy make-up

What will happen to me in the study?

Before you begin the experiment, you will be fitted with a special cap that holds a large number of electrodes against your scalp. These electrode sensors will detect any changes in electrical brain activity (the electroencephalogram, or EEG) that is related to the task you will perform. A small number of loose electrodes will also be stuck onto your face and head with sticky tape to measure blinking and eye movements.

It is necessary to clean the area under each electrode site to ensure good electrode-skin contact. This may cause mild discomfort (due to the abrasiveness of the cleaner), and sometimes causes a little redness. To ensure good contact between the skin and all the electrodes, it is necessary to apply conductive, soluble gel which is best removed by washing your hair. The procedure should not be painful and is very safe. The application of electrodes usually takes around 30-45 minutes, but this may vary dependent on hair thickness.

Once the EEG cap has been attached, you will perform a task that involves memorising words that will appear on a computer screen. Before each task you will be told about what rewards you will receive for later remembering the words. Next, your memory for the words will be tested. After the study you will receive 10/20 % of your winnings. The maximum amount you can win is [2X 14 g bag of chocolate buttons]. You will be given full instructions and get an opportunity to practice some elements of the task before the real experiment begins. After the study you may also be asked to complete some short questionnaires. These tasks will take around 1hr 15 and the complete session will take around 2 ½ - 3 hours, including equipment set-up. You will have the opportunity to take breaks during the session.

What will happen if I don't want to carry on with the study?

We want to emphasise that you may withdraw from the study at any time, and if you refuse to take part or decide to withdraw you will not suffer any penalty or loss of rights. If you do choose to withdraw or are no longer able to participate, then unless verbally directed otherwise the study investigators will keep the data collected up to that point. Your participation can be withdrawn by the study investigators.

Are there any negative side-effects/risks?

We do not anticipate any significant side-effects although some participants may feel a little tired after the test session. There are some risks to participants if they are intolerant or allergic to milk chocolate without the participant's prior knowledge of this allergy; however you will not be required to consume any chocolate during the study itself.

What will happen to the results of the research study?

The results of this study will be used to better understand the brain processes that support word memory and how they are affected by rewards. The results of the study may be published for scientific purposes, but your records or identity will not be revealed unless required by law.

What happens to the information I provide?

Participation in this study guarantees confidentiality of the information you provide. No one apart from members of the research team and carefully selected research collaborators will have any access to the information you provide. Your name and any other identifying information will be stored separately from your data in a securely locked filing cabinet. Questionnaires will be stored in a securely locked room for as long as is required by the Data Protection Act, and then they will be destroyed by our confidential shredding service. You may withdraw from the study at any time, and if you refuse to take part or if you decide to withdraw, you will not suffer any penalty or loss of rights. Your participation in the study can be withdrawn by the study investigators.

What if there is a problem?

If you have a concern about any aspect of the study then you should speak with Dr Zara Bergström, who is director of the study. She can be reached on 01227 827507. If you remain unhappy and wish to complain formally, you can do this through the School of Psychology Chair of Ethics. Further details can be obtained from the School of Psychology General Office on 01227 824775.

Who is organising the research?

The research is organised by Louisa Salhi, a PhD student, under the supervisor of Dr Zara Bergström who works within the School of Psychology.

Who has reviewed this study?

The study has been approved by the School of Psychology, University of Kent Research Ethics committee.

If you would like a copy of the consent form or information sheet to keep, please ask the researcher. If you have any complaints or concerns about this research, you can direct these, in writing, to the Chair of the Psychology Research Ethics Committee by email at: psychethics@kent.ac.uk.

Appendix 4.2.1.2: Experiment 4– EEG Inclusion Criteria Questionnaire

Inclusion criteria Checklist?

Are you aged between 18 and 35? **Yes / No**

Is English as your Native language? **Yes / No**

Are you right handed? **Yes / No**

Do you have normal or corrected to normal vision with glasses/ contacts? **Yes / No**

Do you have hair braids/ weave/ hair extensions? (If so this will not allow us to fit the EEG cap effectively). **Yes / No**

Is your hair type very thick/ very curly? (If so this will not allow us to fit the EEG cap effectively, please talk to the experimenter about your hair type if you are unsure). **Yes / No**

Do you have Dyslexia? **Yes / No**

Are you neurologically healthy? **Yes / No**

Are you taking any psychoactive medication such as anti-depressants? **Yes / No**

Are you intolerant to milk chocolate? (If so let the experimenter know before hand and we can provide dark chocolate instead of milk) **Yes / No**

Appendix 4.2.1.3: Experiment 4– Consent Form

RESEARCH INFORMED CONSENT FORM

Title of Project:	Electrical brain activity associated with word retrieval.	**Ethics Approval Number:**	201614805240054138
Investigator(s):	Louisa Salhi Dr Zara Bergström	**Researcher Email:**	lmas6@kent.ac.uk

Please read the following statements and, if you agree, initial the corresponding box to confirm agreement:
Signatures:

	Initials
I confirm that I have read and understand the information sheet for the above study. I have had the opportunity to consider the information, ask questions about the risks involved and have had these answered satisfactorily.	________
I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason.	________
I understand that my personal information and data will be treated as strictly confidential and will not be made public or shared with any other person outside of the research team.	________
I understand that my data may be published in a research article, but that no personal details will be divulged and that it will not be possible to identify any responses as my own.	
I confirm that I am over 18 years of age and that I freely agree to participate in this study.	

______________________	______________________	______________________
Name of participant (block capitals)	Date	Signature
______________________	______________________	______________________
Researcher (block capitals)	Date	Signature

If you would like a copy of this consent form to keep, please ask the researcher. If you have any complaints or concerns about this research, you can direct these, in writing, to the Chair of the Psychology Research Ethics Committee by email at: psychethics@kent.ac.uk. Alternatively, you can contact us by post at: Ethics Committee Chair, School of Psychology, University of Kent, Canterbury, CT2 7NP.

Appendix 4.2.1.4: Experiment 4– Instructions

WELCOME TO THE CHOCOLATE MEMORY STUDY

During this study you will get chocolate buttons for correctly remembering lists of words.

PRESS SPACE BAR TO CONTINUE

////

Study instructions

Please try to memorize the list of words, all the words will be used in the test.

Every time you correctly recognise a word later you will win a chocolate button.

If you have any questions please ask.

////

Visual search task instructions

During the next few trials you will see coloured pictures presented on the screen.

We want you to look a WHITE TRIANGLE. Not all the pictures will have a white triangle.

Is there a WHITE TRIANGLE on the screen?

Use the Keyboard to respond by pressing:

YES there is a WHITE TRIANGLE, press the 'x' button

NO there is NOT a WHITE TRIANGLE, press the 'n' button

////

Try to answer as quickly and accurately as possible.

If you have any questions please ask the experimenter, if not press ‘space bar’ to start.

Test instructions

Firstly , Is the word old?

Use the Keyboard to respond

'Z' Key -> 'YES': it is an old word from the study list

'M' Key-> 'NO': it is a new word

////

Please hold down the keys to increase your confidence rating.

The longer you hold the key down, the more confident you are. This will show on the screen as a colour change from bright GREEN to bright RED.

'GREEN'-> 'unsure'

'RED' -> 'sure'

////

Try to respond as accurately as possible, every time you correctly recognise a word as an old or a new word you will win a chocolate button.

Please feel free to ask any questions.

////

End of Tests Screen

Please tell the experimenter that you have reached this stage.

They will instruct you about what todo next.

////

Surprise Test Instructions

Now you are about to start the practise trials for the final test.

In this part of the experiment you will see a word presented in the middle of the screen.

You have seen some of these words before either in the;

-> STUDY and TEST phases

-> TEST phase ONLY

-> NEW

Please respond as accurately as you can to the questions.

////

Firstly, have you seen the word before?

Z -> Yes I have seen the word before

M-> No, it is a new word

Hold down the keys to increase your confidence rating.

////

If you have seen the word before, we then want you to think about where you saw the word.

You will be asked two questions;

-> Did you see the word in the Study phase?

-> Did you see the word in the Test phase?

For each question we would like you to respond either Yes or No and indicate your confidence in this decision.

It is possible that the word appeared in both stages, therefore if you remember that this was the case, respond yes to both questions along with your confidence in this judgment.

Z -> Yes

M-> No

Hold down the keys to increase your confidence rating.

////

Please as the experimenter is you have any questions, if not please start the practise trials.

When you are ready to start press SPACE BAR.

////

Now you are about to start the final test.

In this part of the experiment you will see a word presented in the middle of the screen.

You will be asked is you have seen some of these words before and where you saw them.

You have seen some of these words before either in the;

-> STUDY and TEST phases

-> TEST phase ONLY

-> NEW

Please respond as accurately as you can to the questions. Remember to give your confidence response.

Let the experimenter know when you're ready to start.

Appendix 4.2.1.5: Experiment 4 - Post Experiment Questionnaire

Strategy Questionnaire- *Word Memory.*

1: Did you use any techniques or strategies to help you **learn** the words in the learning tasks? If so, please briefly tell us about these techniques. *(Please circle and write in the space below)*

YES NO

2: How much **effort** did you put into learning the words in the **learning tasks**?*(Please circle and write in the space below)*

(no effort) <- 1 2 3 4 5 -> (high effort)

Comments:

3: Did you use any **techniques or strategies** to help you recognised the words in the **recognition test**? If so, please briefly tell us about these techniques. *(Please circle and write in the space below)*

YES NO

Comments:

4: How **motivated** were you to win the maximum amount of chocolate during the **recognition test?** *(Please circle and write in the space below)*

(not motivated) <- 1 2 3 4 5 -> (highly motivated)

Comments:

5: How **difficult** did you feel like the **test** was to **recognise** the words you had earlier tried to learn? *(Please circle and write in the space below)*

(not difficult) <- 1 2 3 4 5 -> (very difficult)

Comments:

6: How much **effort** did you put into recognising the words in the **recognition test**?*(Please circle and write in the space below)*

(no effort) <- 1 2 3 4 5 -> (high effort)

Comments:

7: Did you find that certain words were easier to recognise than others in the **recognition test**? If so, please briefly tell us about these words and why you through they were easier?. *(Please circle and write in the space below)*

YES NO

Comments:____.

8: Did you **anticipate** the **final surprise test** at the end? *(Please circle)*

YES NO

Comments:____.

9: How much **effort** did you put into recognising the words in the **final surprise test**?*(Please circle and write in the space below)*

(no effort) <- 1 2 3 4 5 -> (high effort)

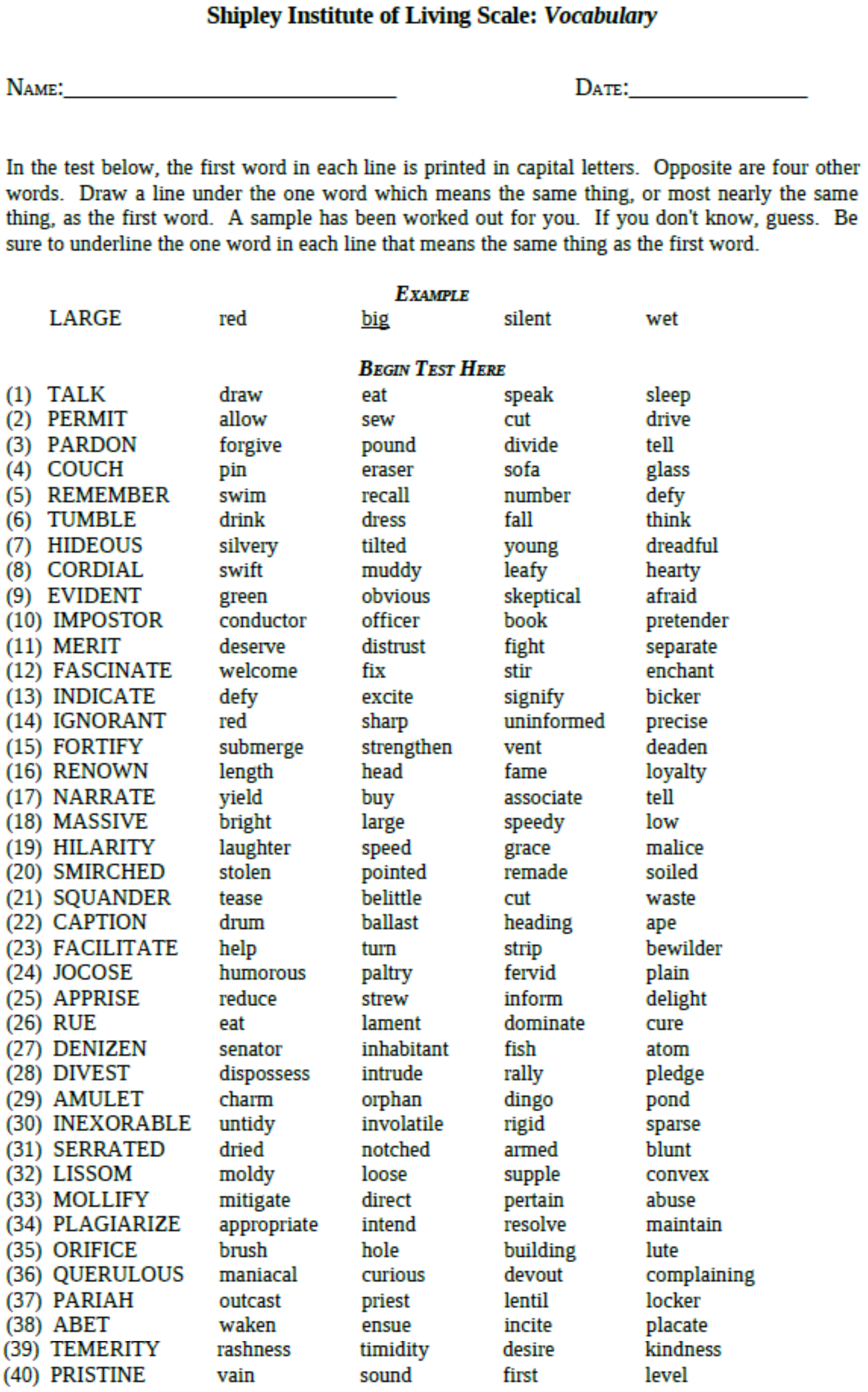
Comments:____.

10: Did you find that certain words were easier to recognise than others in the **final surprise test**? If so, please briefly tell us about these words and why you through they were easier?. *(Please circle and write in the space below)*

YES NO

Comments:____.

Appendix 4.2.1.6: Experiment 4- Shipley Vocabulary Test



Shipley Institute of Living Scale: *Vocabulary*

NAME:______________________________ DATE:________________

In the test below, the first word in each line is printed in capital letters. Opposite are four other words. Draw a line under the one word which means the same thing, or most nearly the same thing, as the first word. A sample has been worked out for you. If you don't know, guess. Be sure to underline the one word in each line that means the same thing as the first word.

EXAMPLE

LARGE	red	big	silent	wet

BEGIN TEST HERE

(1) TALK	draw	eat	speak	sleep
(2) PERMIT	allow	sew	cut	drive
(3) PARDON	forgive	pound	divide	tell
(4) COUCH	pin	eraser	sofa	glass
(5) REMEMBER	swim	recall	number	defy
(6) TUMBLE	drink	dress	fall	think
(7) HIDEOUS	silvery	tilted	young	dreadful
(8) CORDIAL	swift	muddy	leafy	hearty
(9) EVIDENT	green	obvious	skeptical	afraid
(10) IMPOSTOR	conductor	officer	book	pretender
(11) MERIT	deserve	distrust	fight	separate
(12) FASCINATE	welcome	fix	stir	enchant
(13) INDICATE	defy	excite	signify	bicker
(14) IGNORANT	red	sharp	uninformed	precise
(15) FORTIFY	submerge	strengthen	vent	deaden
(16) RENOWN	length	head	fame	loyalty
(17) NARRATE	yield	buy	associate	tell
(18) MASSIVE	bright	large	speedy	low
(19) HILARITY	laughter	speed	grace	malice
(20) SMIRCHED	stolen	pointed	remade	soiled
(21) SQUANDER	tease	belittle	cut	waste
(22) CAPTION	drum	ballast	heading	ape
(23) FACILITATE	help	turn	strip	bewilder
(24) JOCOSE	humorous	paltry	fervid	plain
(25) APPRISE	reduce	strew	inform	delight
(26) RUE	eat	lament	dominate	cure
(27) DENIZEN	senator	inhabitant	fish	atom
(28) DIVEST	dispossess	intrude	rally	pledge
(29) AMULET	charm	orphan	dingo	pond
(30) INEXORABLE	untidy	involatile	rigid	sparse
(31) SERRATED	dried	notched	armed	blunt
(32) LISSOM	moldy	loose	supple	convex
(33) MOLLIFY	mitigate	direct	pertain	abuse
(34) PLAGIARIZE	appropriate	intend	resolve	maintain
(35) ORIFICE	brush	hole	building	lute
(36) QUERULOUS	maniacal	curious	devout	complaining
(37) PARIAH	outcast	priest	lentil	locker
(38) ABET	waken	ensue	incite	placate
(39) TEMERITY	rashness	timidity	desire	kindness
(40) PRISTINE	vain	sound	first	level

Appendix 4.2.1.7: Experiment 4 - Debrief

DEBRIEF FOR PARTICIPANTS.

Study: Electrical brain activity associated with word retrieval.

This study investigated electrical brain activity associated with encoding and retrieval of words and how unintentional encoding of words may vary depending on retrieval success. A constant reward level was maintained throughout to maintain motivation during the task yet was not an experimental factor.

We will look at the amount of words you correctly recognised during the initial test phases, as well as how many new words you recognised in the final surprise test. We asked you about how confident you were in the initial and surprise tests. In the surprise test we also asked you if you could remember where you say the words that you thought were old (source questions), for example you may have been confident that the word was presented in the initial test phase, but unsure if it was presented in the study phase. We will be looking at these judgements in relation to unintentional learning based on prior retrieval accuracy during the first test and whether accuracy on the first test affected firstly surprise test accuracy and source question accuracy.

We will examine the electrical brain activity associated with memory processes linked with intentional learning of words (when you were trying to memorise the materials because you knew your memory would be tested) during the study phases. Whereas during the test phases we will examine the brain activity associated with retrieval processes. We will also examine if being incorrect or correct (your retrieval accuracy) during the initial test affects later recognition during the surprise recognition test. Within the final surprise test we will examine electrical brain activity associated to memory for the non-target, new words that were shown in the first recognition tests. This surprise test examines the effects of retrieval accuracy on unintentional learning during the initial test (when you learned new words during the first test without trying to do so).

We anticipate a difference in brain electrical activity between the correct and incorrect retrieval during the initial test, which will tell us about the brain mechanisms that are important for unintentional learning and remembering. We hope that this study will help us understand memory processes more and how learning and retrieval processes interact, which could in the future lead to useful techniques for improving learning in educational settings or for people with memory problems.

Thank you very much for your important contribution to our research.

Please contact Louisa Salhi at the following email address; lmas6@kent.ac.uk or Dr Zara Bergstrom; Z.M.Bergstrom@kent.ac.uk if you have any questions regarding this study. Similarly if you wish to withdraw your information and contribution to the study please email with your personal created pin to withdraw the data. All of your information and data will be securely and anonymously stored.

Appendix 4.2.2: EEG study- Supplementary analysis

Appendix 4.2.2.1: Pre-EEG trial criteria Behavioural results

As stricter criteria have been applied to the trial selection of the EEG data compared to prior behavioural results, here I test for critical main effects within the full eligible dataset. Constantly with the prior chapter and behavioural pilot study, trials were included if there were more than 10 trials in that condition. Overall, in this analysis there were 39 Younger adults and 40 Older adults included.

Recognition Test 1

On recognition test 1 there were no differences in Discrimination ability (*Pr*; *t*((77)=1.61, *p*=.111, *d*= 0.36; BF_{10}=1/1.395) or bias (*Br*; *t*((77)=1.47, *p*=.147, *d*= 0.33; BF_{10}=1/1.692) across the two age groups as seen in Table A.4.2..1.

Table A.4.2.1. *Discrimination ability (Pr) and Response Bias (Br) during Recognition Test 1, split by Age group.*

Age Group	Discrimination Ability (*Pr*)	Response Bias (*Br*)
Younger Adults	.42 (.17)	.55 (.14)
Older Adults	.48 (.15)	.50 (.17)

Note. Values represent means and SD in parenesis.

Surprise Foil Test Phase

The key results that I wanted to check for replicability were the effects of False Alarms being recognised more on the Surprise test than Correct Rejections, with Slow CRs being recognised lease overall (Table A.4.2.2.). As this had not been examined in Older Adults before it is important to check to see how these effects compared in the Older adult group compared to the Younger adult group.

Table 5. A.4.2.2: *Discrimination ability (Pr) during the Surprise Recognition Test, split by Age group and by Prior Foil condition.*

Measure	Age Group	Prior Foil Condition					
		FA		Fast CR		Slow CR	
Surprise Test Discrimination (*Pr*)	Younger Adults	0.58	(0.15)	0.57	(0.17)	0.55	(0.17)
	Older Adults	0.58	(0.12)	0.51	(0.10)	0.50	(0.10)

Note. Values represent means and SD in parenesis.

Pairwise mixed ANOVA's assessed the difference in discrimination ability (*Pr*) on the surprise test between prior foil conditions between Age Groups as seen in Table A.4.2.3.

Table A.4.2.3. *The results of Mixed 2x2ANOVAs, assessing effects of Prior Foil Condition (FA/ fast CR/ slow CR/ high conf CR/ low conf CR) x Age Group (Younger/ Older Adults) on Discrimination ability (Pr) on the Surprise Test.*

Measure	Prior Foil Conditions	Effect	*F*	*p*	*np2*
Surprise Test Discrimination (*Pr*)	FA x Fast CR	Word Type	**32.97**	**<.001**	**.300**
		Age	**5.33**	**.024**	**.065**
		Word Type *Age	**17.47**	**<.001**	**.185**
	FA x Slow CR	Word Type	**51.89**	**<.001**	**.403**
		Age	**6.18**	**.015**	**.074**
		Word Type *Age	**17.24**	**<.001**	**.183**
	Fast CR x Slow CR	Word Type	**6.97**	**.010**	**.083**
		Age	**12.15**	**.001**	**.136**
		Word Type *Age	.358	.551	.005

Note. Significant values below .05 are shown in bold. N= 79

There was a significant difference in Surprise Test Discrimination between prior FAs and prior Fast CRs with prior FAs being correctly discriminated significantly more than prior Fast CRs (*EMM*; prior FA=.57, *SE*=.015, prior f.CR=.51, *SE*=.017). Younger Adults showed better discrimination overall compared to Older Adults (*EMM*; Younger=.58, *SE*=.022, Older=.50, *SE*=.022). A significant interaction was seen. Within group analysis revealed that Younger Adults did not show a difference in discrimination ability (*Pr*) between prior FAs and prior Fast CRs ($t(38)$= 1.23, p=.227, d= 0.09; BF_{10}=1/2.88) but Older Adults did, with better discrimination for Prior FAs ($t((39)$=6.45, p<.001, d= 0.74; BF_{10}=121292.07).

There was a significant difference in Surprise Test Discrimination between prior FAs and prior Slow CRs with prior FAs being correctly discriminated significantly more than prior Slow CRs (EMM; prior FA=.57, SE=.015, prior s.CR=.49, SE=.017). Younger Adults showed better discrimination overall compared to Older Adults (EMM; Younger=.57, SE=.022, Older=.49, SE=.021). A significant interaction was seen, hen examining this within groups, there was a main effect within both the Younger Adults ($t(38)$= 2.61, p=.013, d=0.20; BF_{10}=3.28) and Older adults ($t((39)$=7.05, p<.001, d=0.92; BF_{10}=724746.36), but to different strengths as Older adults showed a larger difference.

Finally, there was a significant difference in Surprise Test Discrimination between prior Fast CRs and prior Slow CRs with prior Fast CRs being correctly discriminated significantly more than prior Slow CRs (EMM; prior f.CR=.51, SE=.017, prior s.CR=.49, SE=.017). Younger Adults showed better discrimination overall compared to Older Adults (EMM; Younger=.56, SE=.023, Older=.45, SE=.023). No interaction between Foil type and Age group was seen, however for consistency and as we were interested in within group performance, within group analysis was conducted. No differences were seen within Younger Adults ($t(38)$= 1.49, p=.144, d= 0.10; BF_{10}=1/2.09) between foil conditions, but Older adults showed a weak but significant difference ($t((39)$=2.22, p=.032, d= 0.22; BF_{10}=1.53), with Fast CRs being recognised more than Slow CRs.

Source Judgments to the Study Question

Critical prior difference in Source Question responses were seen for the ‘Study’ question, with more errors being made for prior FAs compared to prior CRs. As this is analysis to check for replicable effects to the prior behavioural findings (Chapter 5 Pilot Study), the effects are only shown here for prior FAs and prior CRs. Full analysis is conducted for the EEG sample only.

The Proportions of 'Yes' responses for the 'Study' Question are shown in Table A.4.2.4.

Table A.4.2.4: *Proportion of 'Yes' Responses to the Study Source Questions split by prior Foil type and split by Age Group (Older / Younger adults).*

Source Question	Age Group	Prior Foil Conditions	
		FA	CR
Source Question: Study	Younger Adults	0.57 (0.18)	0.29 (0.23)
	Older Adults	0.44 (0.23)	0.25 (0.23)

Note: Values shown are Means (SD).

A mixed ANOVA with prior Foil Type and Age group indicated a main effect of prior Foil condition on responding 'Yes' to the Source Question 'Study?'. As seen in Table A.4.2.4, prior FAs yielded more incorrect 'yes' responses compared to prior CRs (*F*(1,77)=161.21, *p*<.001, *np2*=.677; *EMM*: prior FA=.50 *SD*=.02 , prior CRs=.27, *SD*=.03). There was no main effect of Age group (*F*(1,77)=5.10, *p*=.027, *np2*=.062; *EMM*: Younger Adults=.43 *SD*=.03 , Older Adults=.34, *SD*=.03). There was however a significant interaction; within group analysis reveal that there was a stronger effect within the Younger adults (*t*(38)= 8.65, *p*<.001, *d*= 1.33; BF_{10}=6.415^{e+07}), compared to the Older adults (*t*(39)= 10.14, *p*<.001, *d*=0.86; BF_{10}=4.850^{e+09}), but both showed the same direction of responses.

Appendix 4.2.2.2: EEG Behavioural Dataset – Additional analysis

Recognition Test 1

The Foil conditions splits have yielded sufficient differences in Correct Rejections to adequately compare to the False Alarms in relation to Confidence measures on Test 1 (Table A.4.2.5). The median splits created another two foil conditions based on splitting Correct Rejections on Confidence Scores (High Confident (hc.CR) and Low Confident (lc.CR) Correct Rejections).

Table A.4.2.5. *Response times and Confidence measures on the first test, shown separately for Foil conditions.*

Measure	Age Group	FA		High Conf CR		Low Conf CR	
Response Time (MS)	Younger Adults	591	(329)	397	(188)	681	(368)
	Older Adults	702	(639)	382	(309)	721	(585)
Confidence (MS)	Younger Adults	742	(392)	1403	(464)	534	(392)
	Older Adults	602	(497)	1343	(776)	407	(326)

Note: Values are means (*SD*) of subject level Medians for Response Times and Confidence.

Surprise Foil Test Phase

It was predicted that Highly Confident CRs (hc.CR) would be recognised more than Low Confident CRs (lc.CR), with FAs being recognised most of all. Discrimination scores (*Pr*) are shown in table A.4.2.6 along with associated Confidence in that subsequent recognition decision.

Table A.4.2.6 *Discrimination ability on the Surprise Test shown in the table split by prior Foil Condition. Confidence scores (ms) also shown split by prior Foil condition.*

Measure	Age Group	Prior Foil Condition					
		FA		High Conf CR		Low Conf CR	
Surprise Test Discrimination (*Pr*)	Younger Adults	0.60	(0.14)	0.60	(0.15)	0.60	(0.14)
	Older Adults	0.58	(0.12)	0.51	(0.10)	0.50	(0.10)
Confidence (*ms*)	Younger Adults	1440	(317)	1486	(321)	1432	(330)
	Older Adults	1220	(833)	1233	(897)	1186	(903)

Note: Values shown are Means (SD). Confidence shown in *ms.*

Separate mixed ANOVA's assessed the difference in discrimination ability (*Pr*) on the surprise test for the Prior Foil Conditions (prior FA, prior hc.CR, prior lc.CR) between the Age Groups (Younger Adults, Older Adults) have been conducted in a pairwise manner in association to pairwise predictions (Table A.4.2.7).

Table A.4.2.7. *The results of Mixed 2x2ANOVAs, assessing effects of Prior Foil Condition (FA/ fast CR/ slow CR/ high conf CR/ low conf CR) x Age Group (Younger/ Older Adults) on Discrimination ability (Pr) on the Surprise Test.*

Measure	Prior Foil Conditions	Effect	*F*	*p*	*np2*
Surprise Test Discrimination (*Pr*)	FA x High Conf CR	Word Type	**19.12**	**<.001**	**.24**
		Age	3.44	.068	.05
		Word Type *Age	**21.78**	**<.001**	**.27**
	FA x Low Conf CR	Word Type	**26.98**	**<.001**	**.31**
		Age	3.86	.054	.06
		Word Type *Age	**21.27**	**<.001**	**.31**
	High conf CR x Low Conf CR	Word Type	<.01	.996	<.01
		Age	**11.61**	**.001**	**.16**
		Word Type *Age	3.77	.057	.06

Note. Significant values below .05 are shown in bold. N= 62

Difference in Surprise Test Discrimination for prior False Alarms and for Correct Rejections split by Test 1 Confidence.

The mixed ANOVA (Table A.4.2.7) with the within subjects' factor of Foil Type (prior FA x prior hc.CR) and Age Group (Younger vs. Older) showed a significant main effect of Foil Type, with surprise test *Pr* being higher for prior FAs than for prior High Confident CRs (*EMM*; prior FA=.59, *SE*=.016, prior hc.CR=.56, *SE*=.016). No significant differences were seen across Age Groups, but a significant interaction was observed. Younger Adults did not differ in their discrimination ability (*t*(31)= 0.26, *p*=.801, *d*=0.02; BF_{10}=1/5.14) however, Older Adults were significantly better at distinguishing prior FAs than prior High Confident CRs (*t*((29)=5.45, *p*<.001, *d*= 0.64; BF_{10} =2847.54).

Similarly, when examining the difference between prior FAs and prior Low Confident CRs, prior FAs were correctly discriminated significantly more than prior Low Confident CRs (*EMM*; prior FA=.59, *SE*=.016, prior lc.CR=.55, *SE*=.015). There also was a trend towards Younger Adults showing better discrimination than Older Adults (*EMM*; Younger=.60, *SE*=.021, Older=.54, *SE*=.022). A significant interaction was seen. Within groups analysis revealed that Younger Adults did not show a difference in discrimination ability (*Pr*) between prior FAs and prior Low Confident CRs ($t(31)= 0.48$, $p=.635$, $d= 0.03$; $BF_{10}=1/4.76$) but Older Adults did, with better discrimination for Prior FAs ($t((29)=6.09$, $p<.001$, $d=0.72$; $BF_{10}=14450.37$)

When comparing prior High Confident CRs to prior Low Confident CRs, there was not a significant main effect of Foil Type, but there was a significant between groups difference, with Younger Adults showing better discrimination ability than Older Adults (*EMM*; Younger=.60, *SE*=.021, Older=.50, *SE*=.022). There was a trend towards an interaction, but no significant within group effects were seen.

Difference in Surprise Test Confidence for prior False Alarms and for Correct Rejections split by Test 1 Confidence.

The only significant effect on Confidence seen is that between prior High Confident CRs and prior Low Confident CRs as seen in Table A.4.2.8, with prior High Confident CRs being more confident on the Surprise Test than prior Low Confident CRs (*EMM*; prior hc.CR=1360ms, *SE*=84, prior lc.CR=1309ms, *SE*=85).

Table A.4.2.8. *The results of Mixed 2x2ANOVAs, assessing effects of Prior Foil Condition (FA/ fast CR/ slow CR/ high conf CR/ low conf CR) x Age Group (Younger/ Older Adults) on Confidence on the Surprise Test.*

Measure	Prior Foil Conditions	Effect	*F*	*p*	*np2*
Confidence Ratings (MS)	FA x High Conf CR	Word Type	1.24	.271	.02
		Age	2.15	.148	.04
		Word Type *Age	0.38	.538	<.01
	FA x Low Conf CR	Word Type	0.57	.453	<.01
		Age	2.07	.156	.03
		Word Type *Age	0.22	.642	<.01
	High Conf CR x Low Conf CR	Word Type	**9.71**	**.003**	**.14**
		Age	2.18	.145	.04
		Word Type *Age	0.04	.838	<.01

Note. Significant values below .05 are shown in bold. N= 62

Source Judgments

The Source Judgments made were 'Yes' or 'No' responses to two Questions; "Was the word in the 'Study' phase?" (Study?) And "Was the Word in the 'Test' phase?" (Test?). Responses here investigated into how many times a 'Yes' response was given. For the 'Study?' this indicated incorrect responses. But for the 'Test?' this indicated correct source judgments. Proportions of 'Yes' responses for both Questions are shown in Table A.4.2.9.

Table A.4.2.9 *Proportion of 'Yes' Responses to the two Source Questions split by prior Foil type and split by Age Group (Older / Younger adults).*

Source Question	Age Group	Prior Foil Conditions		
		FA	High Conf CR	Low Conf CR
Source Question: Study	Younger Adults	0.60 (0.19)	0.27 (0.25)	0.36 (0.25)
	Older Adults	0.45 (0.21)	0.24 (0.24)	0.25 (0.22)
Source Question: Test	Younger Adults	0.83 (0.19)	0.86 (0.22)	0.83 (0.23)
	Older Adults	0.80 (0.14)	0.84 (0.20)	0.85 (0.18)

Note: Values shown are Means (SD).

Mixed 2X2 ANOVAs with Foil Type (2 level design) and Age Group (Younger vs. Older) for both the Study Source Question and the Test Source Question are shown in Table A.4.2.10. As seen in the table the differences in Source Questions were seen only for the incorrect responses to the 'Study?' Source Question, with no differences in correct responses to the 'Test?' Source Question.

Table A.4.2.10 *The results of Mixed 2x2ANOVAs assessing the effects of prior Foil type on responding 'Yes' to the Source Questions (about the Study phase And Test phase) x Age Group (Older / Younger adults).*

Measure	Prior Foil Conditions	Effect	*F*	*p*	*np2*
Responding 'Yes' to the Source Question: "*Study?*"	FA x High Conf CR	Word Type	**141.64**	**<.001**	**.70**
		Age	3.10	.083	.05
		Word Type *Age	**6.03**	**.017**	**.09**
	FA x Low Conf CR	Word Type	**102.80**	**<.001**	**.63**
		Age	**6.46**	**.014**	**.10**
		Word Type *Age	0.68	.412	.01
	High Conf CR x Low Conf CR	Word Type	**14.36**	**<.001**	**.19**
		Age	1.54	.220	.03
		Word Type *Age	**8.75**	**.004**	**.13**
Responding 'Yes' to the Source Question: "*Test?*"	FA x High Conf CR	Word Type	1.22	.275	.02
		Age	0.10	.754	<.01
		Word Type *Age	0.17	.682	<.01
	FA x Low Conf CR	Word Type	0.34	.560	.01
		Age	0.01	.906	<.01
		Word Type *Age	0.88	.352	.01
	High Conf CR x Low Conf CR	Word Type	1.10	.298	.02
		Age	<.01	.983	<.01
		Word Type *Age	5.30	.025	.08

Note. Significant values below .05 are shown in bold.

Study Source Responses split by prior Test 1 Confidence

A mixed ANOVA (Table A.4.2.12) with the within subjects' factor of Foil Type (prior FA x prior hc.CR) and the between subjects' factor Age Group (Younger vs. Older). There was a significant main effect of Foil Type on responding 'Yes' to the Study source question with more incorrect 'Yes' responses for prior FAs than prior High Confident CRs (*EMM*; prior FA=.52, *SE*=.025, prior hc.CR=.26, *SE*=.032). No main effect of Age Group was seen, but a significant interaction was present. Following up on this interaction, when comparing responses between groups within each Foil Type, there is a clear difference in 'Yes' responses for prior FAs ($t(60)$=2.92, p=.005, d=0.74; BF_{10} =8.24), with Younger Adults making more incorrect 'Yes' responses for prior FAs. Whereas, there was no difference when looking between groups at 'Yes' responses to prior High Confident CRs (*t*(60)=0.592, *p*=.556, *d*=0.15; BF_{10}=1/3.33).

Similar analysis with prior FAs and prior Low Confident CRs found increased incorrect 'Yes' responses for prior FAs than prior Low Confident CRs (*EMM*; prior FA=.52, *SE*=.025, prior lc.CR=.30, *SE*=.03). Here there was also a main effect of Age Group, with Younger Adults incorrectly responding 'Yes' more overall (*EMM*; Younger=.48, *SE*=.036, Older=.35, *SE*=.037). No interaction was present.

Finally, when comparing prior High Confident CRs to Low Confident CRs there was a significant main effect of Foil Type, with more incorrect 'Yes' responses to prior High Confident CRs (*EMM*; prior hc.CR=.26, *SE*=.032, prior lc.CR=.30, *SE*=.03). No main effect of Age group was seen, however there was a significant interaction. When examining if there were significantly more incorrect 'Yes' responses within each Foil Type (hc.CR x lc.CR) between Age Groups (Older/Younger) there was no difference between groups in the way they responded to High Confident CRs (*t*(60)=0.59, *p*=.556, *d*=0.15, BF_{10}=1/3.33). There similarly was not a significant difference in the way the Age Groups responded to prior Low Confident CRs, but there was a larger mean difference observed (*t*(60)=1.86, *p*=.067, *d*=0.48, BF_{10}=1.10).

Confidence in Study Source Responses

Differences in Confidence were examined for prior Foil Type for both the 'Study?' and the 'Test?' Source Question as seen in Table A.4.2.12. Raw Confidence scores can be seen in Table A.4.2.11.

Table A.4.2.11. *The average confidence scores for both Source Questions split by prior Foil Type and Age Group.*

Source Question	Age Group	Prior Foil Conditions					
		FA		High Conf CR		Low Conf CR	
Source Question: Study	Younger Adults	979	(382)	1184	(413)	1039	(417)
	Older Adults	673	(552)	761	(663)	578	(553)
Source Question: Test	Younger Adults	1353	(382)	1503	(322)	1405	(345)
	Older Adults	938	(817)	983	(830)	951	(778)

Note: Confidence measures in ms and is the Mean (SD) of the subject level Medians.

Table A.4.2.12. *The results of Mixed 2x2ANOVAs assessing the effects of prior Foil type on responding 'Yes' to the Source Questions (about the Study phase And Test phase) x Age Group (Older / Younger adults).*

Measure	Prior Foil Conditions	Effect	*F*	*p*	*np2*
Confidence in Source Question: *Study?*	FA x High Conf CR	Word Type	**9.34**	**.003**	**.14**
		Age	**9.12**	**.004**	**.13**
		Word Type *Age	1.47	.231	.02
	FA x Low Conf CR	Word Type	0.19	.666	.06
		Age	**11.15**	**.001**	**.16**
		Word Type *Age	3.66	.061	.06
	High Conf CR x Low Conf CR	Word Type	**25.73**	**<.001**	**.30**
		Age	**11.98**	**.001**	**.17**
		Word Type *Age	0.35	.554	.01
Confidence in Source Question: *Test?*	FA x High Conf CR	Word Type	**13.71**	**<.001**	**.19**
		Age	**8.86**	**.004**	**.13**
		Word Type *Age	3.88	.054	.06
	FA x Low Conf CR	Word Type	1.72	.195	.03
		Age	**7.98**	**.006**	**.12**
		Word Type *Age	0.59	.446	.01
	High Conf CR x Low Conf CR	Word Type	**8.56**	**.005**	**.13**
		Age	**10.11**	**.002**	**.14**
		Word Type *Age	2.18	.145	.04

Note. Significant values below .05 are shown in bold.

Confidence in Source Questions split by prior Test 1 Confidence

Study Source Question

A mixed ANOVA with Foil Type (prior FA v prior hc.CR) showed a significant increase in Confidence for prior High confidence CRs than prior FAs on the Study Source Question (*EMM*; prior FA=826, *SE*=60, prior hc.CR=973, *SE*=70). There was a significant difference with Age Group, with Younger Adults being more confident overall (*EMM*; Younger=1082, *SE*=84, Older=717, *SE*=87). No interaction was observed.

When looking between prior Low Confident CRs and prior FAs, there was no main effect of Foil Type, but there was a main effect of Age group, with Younger Adults being more confident (*EMM*; Younger=1009, *SE*=80, Older=625, *SE*=82). Again, no interaction was seen here.

In a separate ANOVA, there was a main effect of Foil Type (prior hc.CR v prior lc.CR) on Confidence, with prior High Confident CRs holding more confidence on the ‘Study?’ Source Question compared to prior Low Confident CRs (*EMM*; prior hc.CR=826, *SE*=60, prior lc.CR=973, *SE*=70). A main effect of Age group was seen again, with higher Confidence ratings overall for Younger compared to Older Adults confident (*EMM*; Younger=1112, *SE*=89, Older=669, *SE*=92). No interaction was seen here.

Test Source Question

A mixed ANOVA with Foil Type (prior FA v prior hc.CR) showed a significant increase in Confidence for prior High confidence CRs than prior FAs on the Test Source Question (*EMM*; prior FA=1145, *SE*=80, prior hc.CR=1245, *SE*=79). A main effect of Age Group was seen, with Younger Adults being more Confident overall (*EMM*; Younger=1428, *SE*=109, Older=960, *SE*=113). There was not a significant interaction, but it is worth noting a trend towards an interaction, with the effect of Foil type only being present in the Younger Adults CRs ($t(31)=3.83$, $p=.001$, $d= 0.43$; $BF_{10}=52.67$) and not the Older Adults ($t(29)=1.30$, $p=.203$, $d= 0.06$; $BF_{10}=1/2.40$).

Examining the difference in Confidence on the Test Source responses between prior FAs and prior Low Confident CRs revealed no main effect of Foil Type on Confidence ratings and no interaction with Age Group. There was however a main effect of Age Group on Confidence, with Younger Adults being more confident for both prior FAs and prior Low Confident CRs (*EMM*; Younger=1379, *SE*=107, Older=944, *SE*=111).

Finally, on a separate ANOVA there was a main effect of Foil Type (prior hc.CR v prior lc.CR), with prior Hight Confident CRs being more confidently responded to on the Test Source Question, than prior Low Confident CRs (*EMM*; prior hc.CR=1243, *SE*=79, prior lc.CR=1178, *SE*=76). There was also a main effect of Age Group, with Younger Adults being more Confident overall compared to Older Adults CRs (*EMM*; Younger=1454, *SE*=106, Older=967, *SE*=110). No interaction was seen.