Understanding Context-binding in Visual Episodic Memory

Emily Victoria Madden

Doctor of Philosophy

University of York
Psychology

July 2024

Abstract

We live in a perceptually rich world, encountering vast amounts of information daily. Extensive research has been conducted regarding our ability to form and retrieve detailed memories, and the myriad factors that determine whether an item will be successfully stored in/retrieved from the long-term memory stores. Whilst memory can be broken down into two main types (semantic and episodic), the present thesis focuses on episodic memories (i.e., memories which are temporally grouped together and contain elements of personal relevance). It is within this memory network that two main factors are examined, namely congruency (i.e., does the item belong in the context) and context-binding (i.e., how well has an individual component in a scene been associated into one cohesive memory trace).

The present thesis aimed to address three main research aims. The influence of semantic congruency and context-binding on episodic memory were examined using behavioural techniques (Chapters 2 and 5). The behavioural findings from Chapter 2 were replicated in Chapter 5, and neuroimaging techniques were used to further examine the neuronal mechanisms underlying such processes. Chapter 3 aimed to examine whether the congruency and context effects observed in visual episodic memory operated in the same manner when applied to a verbal episodic memory paradigm. Finally, the relationship of an object to the scene was examined in Chapter 4, testing classical animacy effects to examine how context-binding operates in visual episodic memory. This thesis provides novel contributions to the field not only in terms of its findings, but in the experimental paradigm used to test the research aims. The key theme emerging throughout the experimental results presented is that context can serve both to facilitate and hinder memory, depending on the relationship of the target to its context and depending on task requirements.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to express my deepest gratitude to my primary supervisor, Dr Philip Quinlan, without whom this thesis would not be possible. Philip has provided invaluable generosity with his time from reading many drafts of my work to fostering engaging scientific debates about the findings and implications of the work. He has made me a stronger researcher and I could not have completed this thesis without him. Additional thanks go to Dr Aidan Horner and Professor Gareth Gaskell for their insight and discussions during my TAP meetings.

I would also like to thank Lance Hilliard for his dedication to helping me in the early stages of my PhD. He spent countless hours on Zoom, with a 6-hour time difference, teaching me to build a website from scratch for custom use. Without his help, my first experiments would not have gotten off the ground during the difficult COVID restrictions. My thanks also go to Dr Anna Guttesen for her unwavering support and guidance throughout the PhD and Masters, helping me to develop my analysis scripts and understanding of the complex statistics involved in analysing EEG/PSG data.

My thanks go to my wonderful family here in the UK – particularly Sarah, John and Indie O’Hagan, Jane and Glyn Knowles. Thank you for helping me make the move from America and for all the love and support, particularly during the challenging parts of my degrees. I could not have gotten through my degrees without you and knowing that you were just a short train journey away. A special thank you to Farah Shippam and Vicky Bailey for your continued friendship, fun nights, support, and additional help with my move to the UK. Many thanks goes to Dr Lucy Spencer and to Fiona Lancelotte for your continued friendship and for being there for me through the tough parts of my PhD. Thank you both for always being open to chat and for being a source of support throughout the process. To my friends in the Psychology Department – Trish, Emma, Dan, Kira, Tim, Karisha, and the numerous friends in the department I have made along the way – I could not have done this degree without you all. Thank you for the pub Fridays, game nights, being an open ear to vent to and for your unconditional support throughout this process. To Team GB – Andrew, Sam, Jade, Esther, Jack, Guy, DingJia, Rutvij, Ewa and Diego – thank you for the friendship, wonderful memories, fun nights, many celebrations, and for all the love and

support you have shown me throughout my Masters and PhD. Thank you to my wonderful housemate Alexis for your support and friendship over the years. Thank you for your patience, kindness, the fun outings, and all the support through the ups and downs of my degree. A very special thank you goes to Tiffanie Jones and Jeannie Kienzle – thank you for the countless hours spent on the phone, proofreading my work, listening to my presentations, and for keeping our friendship going across the pond. The support you have given me over the years has been incredible. To my AI Choir Camp family – thank you for believing in me and being a positive source of inspiration and love over the years. To all my friends back in America, thank you for continuing to support me on this journey.

Finally, I would like to dedicate this thesis to my amazing parents. Thank you for all your love, support and motivation throughout my degree. This distance has not been easy, but you have both been cheering me on every step of the way. Thank you for encouraging me to pursue my dreams and for teaching me from a young age to fight for what I am passionate about. It is with all my love and gratitude that I thank you, Mum and Dad, for everything.

Declaration

I declare that the work presented in this thesis is my own original work and has not been submitted for any examination or degree at this or any other University. This research presented in this thesis was carried out under the supervision of Dr Philip Quinlan, Dr Scott Cairney and Dr Karla Evans. All sources are acknowledged as references.

Some of the data collection was carried out by 3rd year undergraduate project students, who I co-supervised. Some of the stimuli creation and all the data from Chapter 4 was collected by Drew Murton and Sarah Mazari. Data collection was also carried out by research assistants. Some of the data from Chapter 3 was collected by my research assistant, Stephen Robinson. Finally, some of the EEG/PSG data collection and sleep scoring (under my supervision) in Chapter 5 was done by my research assistant Zhiyi Chen. Analysis scripts for the EEG data in Chapter 5 were co-written with Dr Anna Guttesen. However, analyses were carried out solely by myself. The experimental website used to run the studies presented in Chapters 2-4 was created in collaboration with Lance Hilliard. All of the reported data analyses, experimental design, creation of stimuli (with the exception of half the stimuli set in Chapter 4), and all other data collection was performed by myself, under supervision.

A list of conference presentations for each chapter is listed below. The asterisk denotes the presenter.

Data from Chapters 2 and 4 were presented together in poster format at the following conference:

Evans, K. K.*, Madden, E. V., & Cairney, S. A. (2022, May). Relationship between object and scene defines the effects of context on episodic memory over time. Presented at the 2022 Vision Sciences Society Conference in St. Pete Beach, Florida.

Data from Chapter 4 was presented in poster format at the following conferences:

Madden, E. V.*, Cairney, S. A., & Evans, K. K. (2021, August). Visual long-term memory for objects determined by context not their animacy. Presented at the 2021 European Conference on Visual Perception online.

Madden, E. V.*, Cairney, S. A., Evans, K. K. (2021, November). Visual long-term memory for objects determined by their type and strength of binding to scene context. Presented at the 2021 Object Perception, Attention, and Memory Conference online.

The preliminary study design for Chapter 5 was presented in poster format at the following conference:

Madden, E. V.*, Cairney, S. A., & Evans, K. K. (2022, May). Understanding the role of theta in episodic memory formation and consolidation. Presented at the 2022 British Association of Cognitive Neuroscience Conference in Birmingham, UK.

1. Understanding Visual Episodic Memory

1.1 Introduction

We encounter vast amounts of information as we navigate a perceptually rich world from day to day. Researchers have long sought to understand how we make sense of all the information we consume. Once we have experienced a stimulus or event, the question remains as to what factors influence whether it will be stored in our memory or forgotten. The way long-term storage of memory traces comes about is complicated and depends on a variety of components and factors.

From the broadest perspective, memory refers to our ability to encode, store and retrieve information. Encoding can be thought of as a change in state whereby an experience is perceived and transformed from something that one experienced into a memory trace. Retrieval then can be thought of as an outcome of conditions being fulfilled (e.g., memory traces becoming available in the face of relevant cues) that leads to increased probability of retrieval at a later time (Tulving, 1983). However, memory traces must be stored after encoding until they are required to be brought to mind at retrieval. This has led researchers across the last several decades to develop a taxonomic system that organises the various types of memories based on how they are stored. For example, the modal model was introduced in 1986 and argued that there were 3 types of memory: sensory memory, short-term memory and long-term memory (LTM; Atkinson and Shiffrin, 1968). However, Squire and Zola-Morgan (1991) later argued that there were only two main types of memory: declarative (e.g., facts/events) and non-declarative (e.g., skills, conditioning, priming). Prior to this, Tulving (1983) made the distinction that declarative memory is composed of episodic and semantic memory. Whilst semantic memory pertains to facts and general knowledge, episodic memory (EM) refers to memory for events that are temporally grouped together (Baddeley, 1982; Tulving, 1983). The current thesis is concerned with exploring the components that influence episodic memory in more detail.

Episodic memories are said to contain an element of personal relevance (Nuxoll & Laird, 2004; Tulving, 1983). Tulving (1983) theorized that EM was composed of 13 “conceptual elements,” outlining the EM encoding and recollection process (see Figure 1.1). He broke encoding and retrieval down into three main components: things that can be observed (e.g., the event itself or the retrieval cue), processes that influence memory (e.g., encoding or conversion) and states (e.g., the original memory trace or a retrieval experience). These components are not linear or exclusive and can influence how the memory trace is stored or transformed.

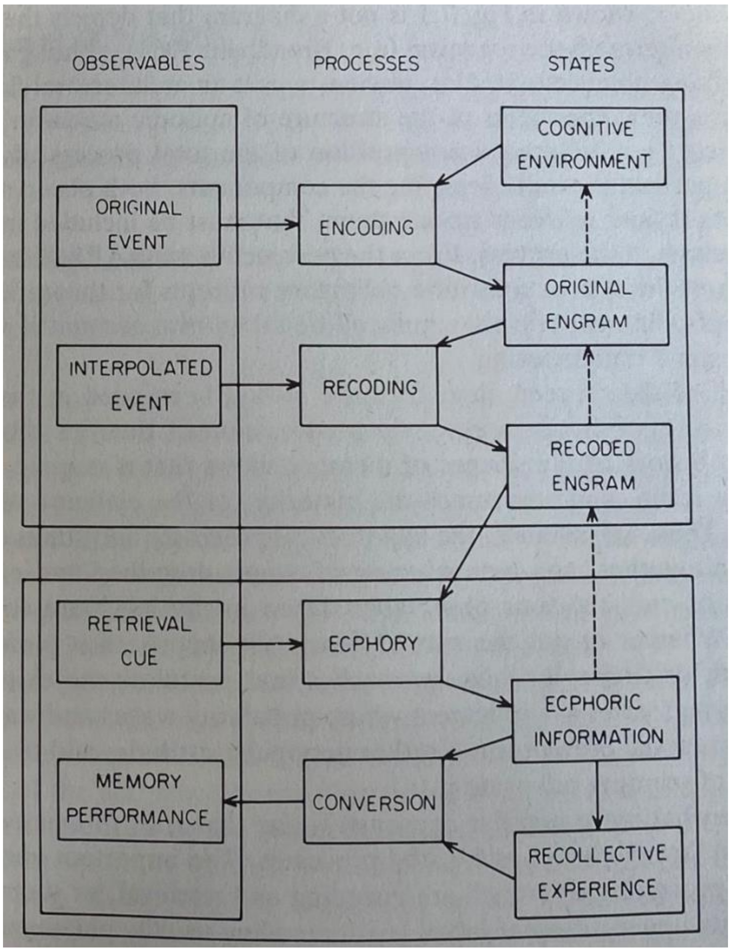

Figure 1.1 Tulving’s (1983) outline of the elements of episodic memory and how they relate to one another (p. 135). At the top of the figure are the elements involved in the encoding of new episodic information. These form the basis for elements involved in retrieval and memory processes (bottom of figure), forming a theoretical system that supports remembering.

More recently, Wang et al. (2012) argued that EM can be thought of in terms of ‘events and episodes.’ They proposed that the ‘events’ pertain to experiences (the ‘what,’ ‘where,’ ‘how’) whereas ‘episodes’ pertain to the temporal details of an event (Wang et al., 2012). This can lead to EM containing an element of personal relevance. However, researchers argue that EM should be considered in terms of the spatio-temporal relationships contained in the memory traces rather than just as a collection of past experiences (Wang et al., 2012). The temporal grouping of spatio-temporal details therefore facilitates retrieval of the memories based on the associations that are formed when the memory trace is encoded (Tibon et al., 2017). The present thesis is therefore interested in the way in which these associations are formed and how they drive episodic memory processes.

The current review will discuss both verbal and visual EM as well as the different components that influence EM. Specifically, the goal is to better understand the way in which EMs are encoded, stored, and retrieved. Of particular interest is the influence of context and decontextualisation on EM, each of which exert a unique influence on how EMs are formed and will be discussed in detail below. Each has been shown to independently influence EM in laboratory experiments, but the ideal laboratory setting is not always translatable to real-world settings. For example, learning material in one context and being tested on it in a different context decreases the amount of information one is able to retrieve. In academic situations students learn in a variety of settings (e.g., in a classroom for lectures, in their rooms revising, etc) and are asked to retrieve the information they have learned in a novel testing environment. If context consistency is the one, key factor, then it would be assumed that performance on tests would be low. However, this is not always the case and students are able to perform well on exams despite the new environment. What is more likely is that the components that have been individually shown to influence EM interact with one another to support efficient and accurate learning. Ultimately, the goal of this thesis is to explore the interactions between the different factors that influence EM in order to further our understanding of the nuanced complexity involved in remembering our highly detailed experiences and events.

1.2 Episodic Memory

1.2.1 Visual Episodic Memory

Early studies were focused on the amount of information that the visual episodic memory store could retain. In an early study, participants were shown 1,000-2,560 colour photographs and their memory for them was tested using a two-alternative forced-choice (2AFC) test after a delay that ranged from 30 minutes to several days (Standing et al., 1970). Memory for the large number of images was high at all time points (> 85%), however the researchers acknowledged that the high retention rates they observed could be due to the overlap in visual cues in the images. In their study they used 300 images containing adult males, all of which shared many physical characteristics. The similarity between the 300 adult males in the images leads to an overlap in visual cues. For example, if several of the images contain men with beards then one could assume that seeing a beard would lead to confusion at a recognition test. However, they observed high memory performance despite the similarity between the stimuli, suggesting that the overlap in visual characteristics might have aided memory for the images in this case. Several years later, Standing (1973) showed participants sets of images that ranged from 20 to 10,000 pictures and tested their memory after a delay of a couple of days. The results showed that EM capacity "follows a power law" based on the number of images presented during the encoding phase (Standing, 1973). Based on the power law assumption, as the number of to-be-remembered items increases so too does the number of items in memory (see Figure 1.2). Standing's power law argument and experimental results suggest that the visual LTM store accounts for increasing memory demands and has a limitless capacity.

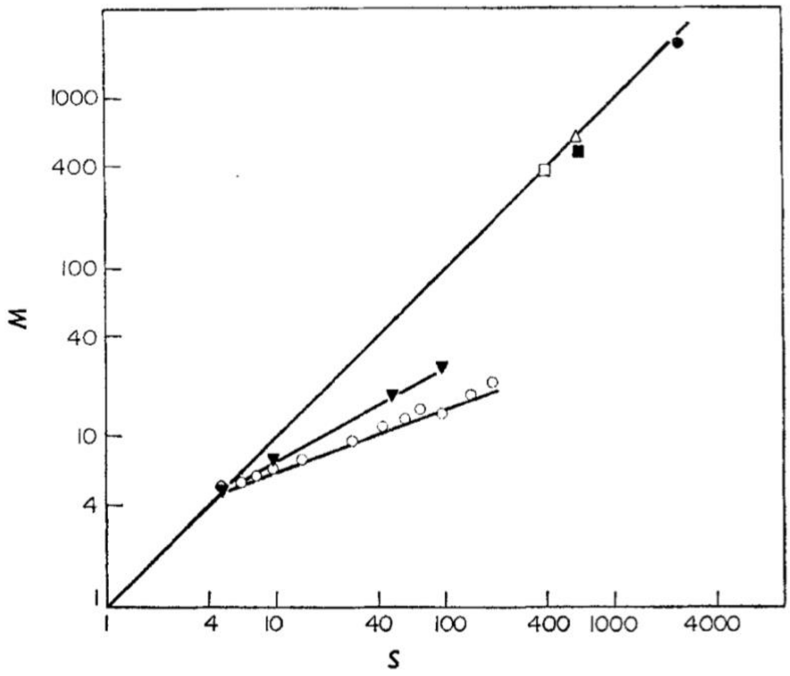

Figure 1.2 Association between memory and set size as a function of the power law across multiple studies. Set size is plotted on the x-axis and items stored in memory is plotted on the y-axis (Standing, 1973). As the number of to-be-remembered items increases, the number of items stored in memory increases in a linear manner.

The observation that visual LTM can store limitless amounts of information has consistently been replicated over the last several decades. For example, Brady et al. (2008) tested the limits at which the visual LTM store can cope without compromising the fidelity of the new memory traces. In their study, participants were shown 2,896 images (2,500 unique images and 396 repeated images). Memory and fidelity for the images was tested using a 2AFC task in which an old image was presented alongside a highly similar foil and participants were asked to indicate which image was the old image. Performance on the memory task was nearly at ceiling with participants correctly identifying the old object between 87.2% (single presentation images) and 92.5% (repeated images) of the time (Brady et al., 2008), despite the large amount of images the participants were asked to commit to memory. They argue that the use of highly similar foils supports the notion that the large capacity for memory does not come at a cost to fidelity since participants were able to correctly identify the old image in the face of a difficult discrimination task.

However, it is not enough to just examine the quantity of information that can be stored; the fidelity of the information stored must also be taken into consideration (Konkle et al., 2010). Konkle and colleagues (2010) further tested capacity and fidelity by presenting participants with 1-16 object exemplars per type of object. Participants studied 2,800 colour images of real-world objects and then performed a 2AFC task to indicate which object had been previously studied. Whilst memory for objects that were the only exemplar for their category was highest (93%), presenting 16 exemplars per category still showed robust memory for the objects (82%). Memory for perceptually similar objects may decrease as the number of similar objects increases, but not to a point that visual LTM for the objects is compromised. Visual EM is able to cope with thousands of similar and richly detailed images without losing the fidelity of the images.

Along these lines, Vogt and Magnussen (2007) examined whether presenting participants with 400 images of doors would hinder LTM due to the high level of overlap in the stimuli. Memory for the doors was tested at four different time points (ranging from 30 minutes to 9 days). Participants completed a 2AFC task in which a target picture was paired with a novel distractor picture. For some of the target images, small details were stripped from the photos (e.g., flowerpots near the door) to assess whether ‘irrelevant’ details aid memory performance. In line with previous work, memory performance across the 9-day period was high, further supporting the large capacity of visual LTM. However, they observed a 20% decrease in performance when small, seemingly irrelevant details were removed from the pictures. They argue that the details present in an image are crucial to the formation of LTMs (Vogt & Magnussen, 2007). One could argue that the entire image is stored in LTM as a holistic representation. Removing ‘irrelevant’ details before presenting an image in a recognition test might decrease the likelihood of correctly recognising the image due to the small decrease in how well the cue and original memory trace match.

More recently, Evans and Baddeley (2018) used the door stimuli from Vogt and Magnussen (2007) to test whether there was an interaction between the amount of detail available in an image and whether or not one is aware that memory for the images will be tested. As expected, participants who were aware that a test would follow performed better than those who were unaware. The doors with more detail were also significantly better remembered than the doors with details stripped such that the removal of details led to an increase in false alarms when presented with a novel door (Evans & Baddeley, 2018). This suggests that the rich details from the original images were encoded and stored as a cohesive memory trace and that stripping the details before presenting the items at test decreases the quantity of elements that can serve as retrieval cues that would facilitate successful and accurate retrieval. It leaves the visual memory system susceptible to errors and incorrectly attributing a new image as having been previously seen, further supporting the notion that visual long-term EM preserves highly detailed representations of each event. One could further argue that these studies (Brady et al., 2008; Evans & Baddeley, 2018; Konkle et al., 2010; Vogt & Magnussen, 2007) highlight the robustness of the visual long-term EM store via its ability to flexibly adapt to the increasing memory demands that one places upon it.

If visual LTM has a limitless capacity that preserves the fidelity of the original images, then what factors might be contributing to this level of capacity and fidelity? Friedman (1979) proposed that there are two key elements present in images used for visual EM tasks: whether or not a to-be-remembered element was present (episodic) and what that element looked like (descriptive). The fact that images can provide a snapshot of previously and/or frequently encountered scenarios (e.g., a kitchen or a living room) gave rise to the argument that memory for the elements of an image, both episodic and descriptive, should be poor regardless of when the test takes place (Friedman, 1979). Friedman (1979) further suggested that using conceptually similar targets and distractors reduces recollection accuracy due to confusion, providing a counter argument to Standing's (1973) attempt to explain their high retention scores. Other researchers have argued that visual stimuli are better remembered than verbal stimuli due to the extra details images elicit (Konkle et al., 2010). For example, the word 'dog' only provides basic-level category information and will bring to mind a variety of dogs. A picture of a dog, however, will provide more detailed information (e.g., a chocolate lab) that narrows down the endless possible matches that the generic word 'dog' elicits (Konkle et al., 2010).

Other researchers have turned to the neural regions that might support the limitless capacity of visual LTM. The hippocampus (HC) has been identified to play a key role in EM, especially during the encoding of EMs (see Berron et al., 2016; Horner et al., 2015; Neunuebel & Knierim, 2014; Norman et al., 2008; Tse et al., 2007). The HC is believed to be responsible for the essential function of binding of elements in episodic events into EMs (Horner et al., 2015). Whilst it is commonly agreed that the HC is critical for encoding, there is less agreement about the mechanisms that determine how the events are encoded and consolidated. For example, the context in which a to-be-remembered item (i.e., the target) is presented in effects how well the target is remembered. Another example is in the effects of semantic congruency on visual EM. Some researchers argue that congruency with a background increases memory for the target whilst others argue that incongruent targets stand out more and are more memorable. The encoding and consolidation of visual episodic memory is nuanced and sensitive to a variety of factors, discussed in detail later in this review.

1.2.2 Verbal Episodic Memory

Some researchers have argued that visual and verbal memory share the same qualitative processes and features. Baddeley (1999) argued that even when presented with visual stimuli, memory processes involve verbalising the details of the event. He suggested that this leads to the event no longer being purely visual. Further, he argues that the same rules that govern learning and retrieval of verbal material are shared with learning and retrieval of visual information (pg. 16), suggesting that two systems share processing characteristics. For example, Standing (1973) suggested that both visual and verbal EM storage systems “follow a power law for capacity, both decline in terms of items correct (but not in terms of the detectability index) when the number of alternatives in the recognition test is increased, both show a comparable decline in performance when a recall task is substituted, and both follow a power law for retrieval time.” Accurate memory for visual (images) and verbal (sentences/words) material is comparable (Shepard, 1967), suggesting that the two modalities might share encoding and consolidation processes.

Verbal episodic memory has been shown to be robust in terms of capacity and retrieval (see DeLong et al., 2005; DeLong & Kutas, 2020; Hovhannisyan et al., 2021; Mak et al., 2021; Mak & Twitchell, 2020; Staresina et al., 2009). Much like visual EM, verbal EM is sensitive to semantic relatedness (DeLong & Kutas, 2020; Kutas & Hillyard, 1980; Mak et al., 2021; Mak & Twitchell, 2020), links to previous experiences and expectations (DeLong et al., 2005; DeLong & Kutas, 2020; Kutas & Federmeier, 2000; Thornhill & Van Petten, 2012) and the amount of information presented (Standing, 1973). Whilst early research focused on comparing the two modalities, modern research has separated the two and studied them independently. In order to better understand episodic memory as a whole, this thesis proposes that the factors that influence visual episodic memory should also be examined in regard to how they influence verbal episodic memory, potentially providing insight into processes that the two modalities share and deepening our understanding of how we make sense of the complex world around us.

1.3 Factors Influencing Episodic Memory

1.3.1 Context

‘Context’ was originally defined by Tulving (1983) as the “setting” or “cognitive” environment surrounding a memory. In terms of EM, it can be used to refer to the elements of an event (e.g., spatial or temporal) that tie the event into a cohesive unit (Yonelinas et al., 2019). For EMs, context can exert a variety of influence over how memories are encoded, consolidated, and retrieved. Some examples of the effect of context on EM are its influence via the environmental context (Cairney et al., 2011; Eich, 1985; Godden & Baddeley, 1975) or an individual's internal context at the time of encoding (Criss & Howard, 2015), and the influence of the semantic context surrounding a target word (Packard et al., 2020; Staresina et al., 2009; Tibon et al., 2017) or object (Davenport & Potter, 2004; Gerver et al., 2020; Hollingworth, 2006). It is not necessary for the participant to be aware that contextual associations are present in order for them to exert an influence on EM (Oliva & Torralba, 2007), suggesting that EM implicitly stores the memory trace as a cohesive and complete event.

Context is thought to provide a framework for the elements of an episode, guiding and facilitating expectations of plausible items appearing in that scene for object identification and retrieval (Hayes et al., 2007). Animal studies show that EMs are initially linked to the context in which they are encountered but that, with the passage of time, context sensitivity diminishes and generalisation into a more gist-like representation occurs (Winocur et al., 2007). For example, Winocur et al. (2007) fear-conditioned rats in different test chambers (i.e., contexts) and tested their memory for the conditioned response after a 1-day or 28-day delay in either the same chamber or in a different chamber. At the shorter delay, rats placed back in the same contextual environment that they were conditioned in resulted in longer freezing times. However, after a 28-day delay the rats spent roughly the same amount of time freezing regardless of whether the contextual environment was the same or different. Wiltgen et al. (2010) used the same contextual fear-conditioning paradigm, also testing memory after 1-day vs 28-days, and replicated the results of Winocur et al. (2007). However, they argue that the ability of the rats to distinguish between old and new contexts at the short time delay supports the notion that the context fear memories are becoming less detailed with the passage of time (Wiltgen et al., 2010). These animal studies provide support for the idea that context effects diminish as the time between learning and test increases, making the memory traces less dependent upon their contexts.

Understanding context as one factor in EM processes has demonstrated that there are many ways in which context exerts an influence on the consolidation and retrieval of EMs. In academic settings, the context that students learn in almost always varies from the context in which they are asked to retrieve the information they have learned. Assuming that the role of context in EM is varied, the question remains as to how one can diminish the potentially detrimental effects of changing contexts in order to maximise student performance. The current work is particularly interested in three of the ways context influences EM: how targets are bound to their contexts, the detrimental effect changing context has on memory and how the bond between an object and its context is weakened through the process of decontextualisation. This section will address these three influencing factors in detail.

1.3.1.1 Context Binding

Since EMs are known to be spatially and temporally grouped together, researchers have turned to the ways in which these memory traces are bound to the contexts in which they occur. The context facilitates the creation of a holistic representation of an event by linking the elements of the event to their episodic context (Hollingworth, 2006). The contextual binding theory posits that when new EM traces are encoded, the HC binds the target to the context it was encountered in, creating the critical link between the two in order to facilitate retrieval later on (Yonelinas et al., 2019). Episodic memory is said to be 'context bound' as the contextual details support accurate retrieval, particularly when the contexts at study and test match (Criss & Howard, 2015). Moreover, the 'context' does not have to be physically present in order to influence how a target is bound to its context: simply instructing an individual to imagine a target item in a vivid context is sufficient for context binding to occur (Eich, 1985).

Eich (1985) had participants learn lists of words in a specific environmental context. For some of the words in the list, they were instructed to create a vivid mental scene in which the object was present. For the remaining words, they were instructed to bring the isolated object to mind. Memory for the lists of objects was tested either in the same room as where they learned the lists or in a new room. As expected, the change in environmental context led to a significant decrease in memory for the objects paired with a mental context. Objects that were associated with a vivid mental image were better remembered overall than ones brought to mind in isolation. More interesting however was the finding that the change in environmental context did not affect the recall of items lacking a mental context. Eich (1985) argued that the integration of a mental context was a crucial requirement of being able to recall the objects from the list. The lack of an environmental context effect on isolated objects could be due to a higher cognitive demand needed to only bring to mind the object and not previously learned contextual associations, leaving little cognitive resources available for the environmental context to be encoded (Eich, 1985). The fact that the change in environmental context led to decreased memory for the integrated mental context items provides early evidence that items are bound and dependent upon all the contextual details available at the point of encoding.

Hollingworth (2006) was also interested in whether items that had previously been associated with a context would be adversely affected by being recalled in isolation. In their study, participants were shown a series of images depicting real-world scenes, containing seven fully visible objects. After each scene was presented, a 2AFC recognition test occurred in which memory for one of the objects was tested. The targets were either unchanged from the original image, the original object but rotated 90° or it was a new object from the same semantic category. The 2AFC test was either with the target in isolation (on a grey background) or with the target in the original scene context. Results showed that objects tested in isolation resulted in decreased recognition accuracy compared to those presented in the original scene. This provides support for the argument that, when presented with real-world scenes, the elements of the scene are stored as a holistic representation and that context affects the targets on an exemplar level (Hollingworth, 2006). Taking an item out of its context leads to decreased recognition memory potentially due to the decrease in semantic-level cues at recognition (Hollingworth, 2006).

More recently and based on the previously mentioned study by Hollingworth (2006), Nakashima and Yokosawa (2011) explored whether object-to-scene binding always occurs when objects are presented as part of a scene. They assumed that, if objects are always bound to a scene context, then that context should always facilitate retrieval despite the type of retrieval task. In their study, participants viewed 60 real-world scenes containing 15 objects. Each scene was presented as part of a flicker display in which the scene was presented for a total of 10.88 seconds (240 ms repeated presentations separated by 100 ms black screen until the total display time was reached). Participants were instructed that they would see flickering images and that their task was to decide if one of the objects present in the scene was the same or different from the one in the flickering display. Consistent with Hollingworth (2006), some of the objects were tested either with or without a background context and were either the same object, the same object but rotated 90° or a new, semantically related object. As expected, participants performed better and more accurately recognised objects that were tested within a scene context as opposed to those being tested on a grey background in isolation. Their first study shows that object-to-scene binding occurs quickly, even when the images are disrupted by a flickering presentation. In their second study, the flickering paradigm included a change detection task such that the images either alternated between the target object and a different but related object, or the image did not change. They failed to observe an object-to-scene binding effect when participants were asked to remember the objects as part of a flicker change detection task. One possible explanation as to why context binding occurs for some tasks and not others could be due to the amount of information that the individual needs to remember (Nakashima & Yokosawa, 2011). For example, knowing that the images were not going to change (Experiment 1) meant that the participant could form a mental ‘chunk’ of the information each time it was presented as a memorisation strategy. This type of strategy results in the object being quickly and tightly bound to the scene, with one caveat being that it is contingent upon how much information the individual is being asked to remember (Nakashima & Yokosawa, 2011).

Therefore, when the task requires large amounts of information to be retained in memory, context binding might offer one faciliatory approach to storing EM traces so that they can be retrieved effectively based on the contextual associations present during the original event. As this thesis is primarily concerned with examining the factors that influence EM, a main focus is to better understand how the binding of a target with its context influences EM. Namely, whether the binding of information enhances memory for elements of the events, especially when the amount of to-be-remembered information is high and the amount of time between encoding and retrieval increases.

1.3.1.2 Context Shift Decrement

If context exerts a strong influence over the retrieval of EMs, then it is to be expected that changes in context will result in reduced memory for elements of the episode. A context shift decrement is said to explain observed decreases in memory when items are presented in a different context at test than the one it was presented in during study (Hayes et al., 2007). Conversely, a context reinstatement effect is observed when the context at study and test match, hence the context has been ‘reinstated’ and facilitates accurate retrieval of the target (Smith et al., 2018). Reinstating the context is believed to help facilitate retrieval, but some argue that it is only helpful when instructions prompt the explicit encoding of the entire event (Hanczakowski et al., 2015). Others have argued that a context shift decrement is more pronounced when the target is part of a visually rich scene (Hayes et al., 2007), suggesting that targets embedded in real-world scenes should be more susceptible to a context shift and reinstatement. Additionally, a context shift decrement occurs when memory demands are high and require large amounts of information to be remembered (Evans & Wolfe, 2022).

Hayes et al. (2007) conducted a series of studies that examined under what conditions a context shift decrement occurs. In the first two studies, participants were shown 240 unique objects in differing background scenes. Memory for the object was tested either with the object in its original scene or with the object on a white background. Participants were asked to make judgements about the price of the objects followed by a surprise memory test (Experiment 1) or were asked to only attend to the object and ignore the background (Experiment 2). Both sets of results showed a significant context shift decrement, with a decrease in memory for objects appearing against a white background after having been studied in a visually rich scene. Interestingly, when they changed the experimental conditions such that objects were learned in a white, blank background and then tested in a visually rich scene (experiment 4), a context shift decrement was still observed. Taken together, these results suggest that the object is automatically tightly bound to the scene in which it is learned regardless of whether an individual is asked to remember the context or not (Hayes et al., 2007).

Evans and Wolfe (2022) also observed a context shift decrement after presenting participants with 100 real-world, indoor scenes. Each image was presented for 3 seconds and participants were asked to only focus on the target object inside the red bounding box in order to alleviate the need to search for target objects within the scene. Participants were explicitly told that the background scene was irrelevant and that a memory test would occur. At test, the 100 target objects were presented to participants in new, real-world indoor backgrounds. They found that changing the background between study and test lead to incredibly poor memory performance, supporting the context shift decrement. Moreover, memory scores for objects in a changed background was also worse than when the objects lacked a background. They argue that when a scene is present during the encoding of a memory trace, it leads to a detrimental effect at retrieval if that background is not consistent with the original event (Evans & Wolfe, 2022). Even when participants are instructed to ignore the context, the entire scene is tightly bound and context-dependent (Evans & Wolfe, 2022).

Changing a context between study and test also decreases the likelihood of being able to accurately distinguish targets from distractors, manifesting as a decrease in memory accuracy at test (Murnane et al., 1999). The automatic encoding of the context helps to differentiate one EM from the next and increases the likelihood that the correct memory trace is retrieved when cued (Murnane et al., 1999). Taken together, it is possible that context serves as an aid to encoding by tightly binding all the elements of a scene together. However, it proves particularly useful when it comes to accurate retrieval at a later point in time. A context shift decrement can therefore disrupt the encoding and retrieval processes, resulting in a decrease in memory accuracy. In terms of the main research goals, the studies presented throughout this thesis examine the conditions in which a context shift decrement emerges and they ways in which it might be overcome in order to minimise its detrimental effect on memory performance.

1.3.1.3 Decontextualisation

If context is incidentally encoded and is an important aspect of accurate retrieval, then the question remains as to whether there is a way to weaken the bonds without sacrificing memory performance. Decontextualisation is the gradual process of stripping detailed EMs in to more abstract, gist-like memory traces. The traces become more gist-like as the associative links between the rich contextual details are weakened with time (Cox et al., 2014; Jurewicz et al., 2016; Piolino et al., 2009; Winocur et al., 2007; Winocur & Moscovitch, 2011). Since the HC is responsible for linking the associative elements of an event, the overwriting of HC representations has been argued to facilitate decontextualisation (Cox et al., 2014). This is due to the sparing of representations (e.g., generalised representations of the target object) that have already been transferred to LTM in the NC (Cox et al., 2014). Traces still in the HC are updated whilst other elements of the event are left untouched. Recognition tasks are suggested to be particularly sensitive to decontextualisation effects (Jurewicz et al., 2016; Piolino et al., 2009) due to the fact that successfully making an old/new judgement does not rely on bringing to mind all of the contextual details associated with a specific event (Mandler, 1980). This is speculated to lead to a "Remember-to-Know" shift due to repeated exposure to highly similar events (Piolino et al., 2009), making recognition memory more sensitive to decontextualisation effects.

Researchers have tried to induce decontextualisation effects by changing the background context each time a target is presented to a participant. For example, if the target object is a table in a dining room, then it would be presented multiple times during the study phase with each presentation being the same table in a new dining room scene. This manipulation exploits the active alteration of the memory trace by changing the associative links and decreasing the contextual dependence at retrieval. It serves as an attempt to speed up the decontextualisation process. For example, Cox et al. (2014) had participants learn 120 words that were presented on landscapes, each presented four times. Words were split into two retrieval sets (60 each session) and memory for the words was tested after a 12-hr delay and a 24-hr delay. Half the cue words (30 each test session) were presented on the same landscape whilst the remaining 30 backgrounds were shuffled and re-paired with novel words. They observed an overall reduction in memory over time, suggesting that context effects are reduced as time passes leading to a decontextualisation of the memory traces (Cox et al., 2014).

Similarly, Jurewicz et al. (2016) examined how episodic memories are decontextualised by pairing words with background videos that served as the contextual cues at retrieval. Participants viewed 365 word-video pairs during the study phase and recognition memory was tested after either a 12-hr or 24-hr delay. At test, they were asked to make old/new judgements pertaining to old words on the same background video (matched), old words with a different background video (non-matched) or new word-video pairs (foils). Whilst they observed a memory benefit for words that appeared in the same background context (reinstatement effect), they failed to observe an effect of decontextualisation after a 12-hr and 24-hr retention period. They propose that more time might be needed in order to facilitate the weakening of the contextual dependency of the episodes (Jurewicz et al., 2016). Whilst there is still a debate about whether or not sleep plays a role in the decontextualisation process (Cox et al., 2014; Jurewicz et al., 2016; Rauchs et al., 2005), it is agreed that time is a necessary factor in how memory traces are decontextualised. Since decontextualisation weakens the associative links between elements of EMs, it offers an alternative explanation as to how context influences the way in which EMs are consolidated for later retrieval. The current research is concerned with exploring the ways in which decontextualisation might offset the detrimental effects of changing contexts in order to increase memory performance long-term and decrease context-dependency.

1.3.2 Consolidation

Consolidation is the process by which fragile memory traces are stabilised and stored in LTM (Hupbach et al., 2007; Nadel et al., 2012; Roesler & McGaugh, 2019; Tambini & D'Esposito, 2020; Winocur & Moscovitch, 2011). Consolidation takes place after an episode has occurred and been encoded. The memory trace is consolidated over time such that the event becomes a 'fixed memory' in the LTM networks (see Figure 1.3 Nadel et al., 2012).

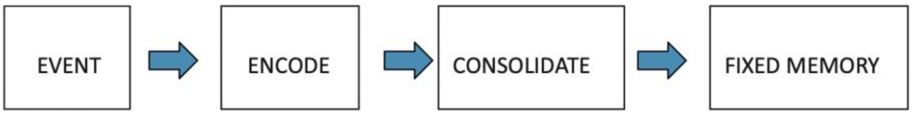

Figure 1.3 Simple pipeline of the encoding and consolidation process (Nadel et al., 2012).

The standard two-stage model of memory proposes that memory is composed of two separate memory stores: a fast-learning, temporary storage site and a slow-learning, long-term store (Born & Wilhelm, 2012; Schreiner & Staudigl, 2020). The standard model of memory posits that consolidation begins once the neocortex (NC) has registered that the HC has created a bound memory trace (Nadel & Moscovitch, 1997). Short-term consolidation immediately binds the elements of the new memory trace together in order for the HC to transfer the memory trace to the NC, reducing the memory's reliance on the HC over time (Nadel & Moscovitch, 1997). It has been speculated that consolidation determines what elements of a memory are stored (Nadel et al., 2012), yet the growing body of research into the consolidation process demonstrates that this is too simplistic. Understanding how neural representations are formed and retrieved has become a central point of research in cognitive neuroscience.

Memory consolidation is typically thought to involve the stabilisation of fragile memory traces into durable memory traces within LTM networks. However, this may occur via the reactivation of the memory trace during quiet periods in which no new information is being encoded (Sutherland, 2000). The HC plays a critical role in reactivation (Nadel et al., 2012; O'Neill et al., 2010; Sutherland, 2000; Tambini & D'Esposito, 2020; van Kesteren et al., 2020), facilitating reactivation brain-wide across modalities (O'Neill et al., 2010) and gradually determining the synaptic distribution allotted to each memory trace before transferring it to LTM (Sutherland, 2000). The HC has the ability to quickly switch between online and offline processing, supporting reactivation during sharp-wave ripples that are not necessarily dependent upon sleep, facilitating the binding of associations in the HC that allow for LTM integration (O'Neill et al., 2010). Further, the reactivation of previously learned information during the encoding of new, related information aids schema development via the integration of associated information into semantically organised memory networks (van Kesteren et al., 2020).

Schemas are defined as knowledge structures that are grouped based on semantic relatedness and are based on past experiences (Brewer & Treyens, 1981). Schemas are believed to facilitate the encoding and retrieval of EMs due to their hierarchical organisation that allows both top-down and bottom-up activation of information when presented with new experiences (Ghosh & Gilboa, 2014). Animals studies show that the creation of a schema is a gradual process (~1 month), but that updating and incorporating new information into an existing schema occurs quickly once the schema has been activated (Tse et al., 2007). Each schema is comprised of multiple episodes that share the same semantic structure (Ghosh & Gilboa, 2014). Therefore, it is essential that schemas remain general if they are going to allow for the integration of new episodes (Ghosh & Gilboa, 2014). Once the schema has been established, consolidation is thought to become more efficient due to the established network of connected NC representations (van Kesteren et al., 2020). The schema serves as a framework that allows new information to be assimilated into the existing networks and allows memory traces to be identified efficiently at retrieval (Anderson, 1984). Therefore, schemas might provide another explanation of the way in which EMs are stored and retrieved. When faced with a partial cue, the schema is activated and facilitates inferential reconstruction of the event (Anderson, 1984). For example, one might remember having seen a kitchen. In the event that they do not remember the specific details of that kitchen, their 'kitchen' schema will be activated, and they can assume that a stove, fridge, sink, etc should have been present. In addition to serving as a retrieval mechanism for EMs, it is likely that schemas also provide a mechanistic account for the storage of EMs. The activation and integration of new EMs into existing schemas is postulated to speed up the processes that lead to the memory trace becoming HC-independent (Tse et al., 2007), providing a glimpse into the structure of how EMs are stored long term.

However, it has been postulated that once memories have been consolidated that they are not intended to remain in a permanent, unmodifiable state (Craig et al., 2021). Consolidated memories have the potential to undergo 'reconsolidation' if they are reactivated by novel experiences, cues or retrieval (Craig et al., 2021; Hupbach et al., 2007; Nadel et al., 2012; Winocur et al., 2009). The reactivation of a memory trace returns it to a labile state, leaving it vulnerable to hippocampal disruption (Winocur et al., 2009) and to being updated or potentially erased when it undergoes reconsolidation (Nadel et al., 2012). Figure 1.4 shows an example course of events once a memory trace is reactivated.

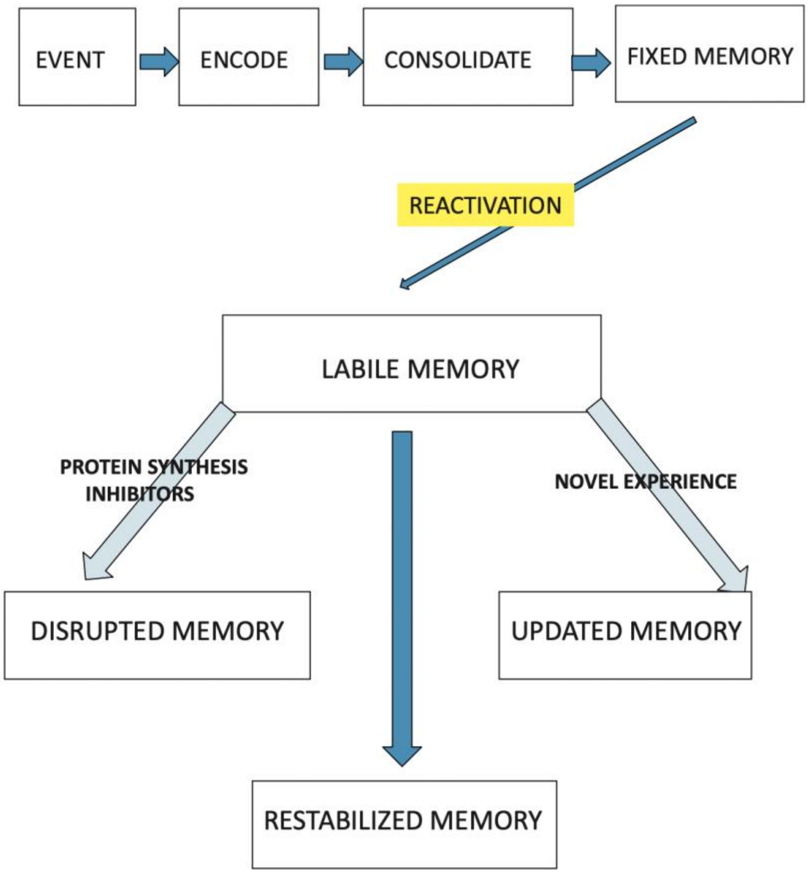

Figure 1.4 Process of reactivation as explained by Nadel et al. (2012). Animal studies show that the injection of protein synthesis inhibitors disrupts the reactivated memory trace. Reactivating the memory trace without any new input leads to the memory trace being restabilised whereas reactivation in response to a novel experience leads to the memory trace being updated (Nadel et al., 2012).

Hupbach and colleagues (2007) conducted two studies in which they examined the way in which reconsolidation facilitates the updating of previously consolidated EMs. In their first experiment, participants were asked to learn a set of physical objects (presented by the experimenter, one at a time from a basket). The next day, one group of participants were asked to learn a new set of objects (no reminder group) whilst the other set were asked “Do you remember this basket and what we did with it” before learning the new set of objects (reminder group). Memory for the objects was tested on the third day. Their results showed that prompting participants to recall what was done with the basket reactivated the objects studied on the first day. This led to a high number of objects being misattributed to the first day of learning compared to the no-reminder group who showed few intrusions of Day 2 objects attributed as Day 1 objects. Experiment 2 used the same experimental paradigm with the only difference being that memory was tested immediately after the second set of objects was learned rather than the next day. Interestingly, the intrusion effect did not replicate when memory was tested immediately after the second learning phase.

Hupbach et al. (2007) argue that their results support a time-dependent process of reconsolidation for EMs, supporting the idea that interference caused by reactivation does not appear immediately. Moreover, the subtle prompt to recall the basket used to hold the objects was sufficient a cue to reactive the original set of objects and return them to a labile state. Reconsolidation provides another approach to the way in which EMs are maintained and consolidated. It leads to the assumption that EMs are 'dynamic' memory traces (Hupbach et al., 2007) that are nuanced in terms of how they are encoded and consolidated.

1.3.3 Time and Forgetting

Researchers agree that consolidation is a slow process and does not happen instantaneously. The transfer of encoded memory traces to their LTM stores takes time, arguably needing weeks to months and even years before stabilisation and consolidation are considered complete (Roesler & McGaugh, 2019). Tulving (1983) argues that one of the most prevalent features of EM is that recollection and memory for the events change over time as manifested in changes in memory performance (pg. 164). Moreover, he argues that the episodic system, in general, is highly susceptible to "transformation and loss of information" (pg. 21), begging the question as to what factors contribute to changes over time and the forgetting of episodic information.

Animal studies have provided evidence of the time-dependent nature of EM consolidation and forgetting. Tse et al. (2007) used mice to investigate the timeframe in which newly learned paired associates are no longer dependent on the HC, indicative of consolidation having taken place. Forgetting was consistently observed when HC lesions were made within 90 minutes of encoding a novel paired associate. However, if the lesions occurred one month after learning had taken place, memory was not impaired (Tse et al., 2007). Similarly, Winocur et al. (2007) fear-conditioned mice in one environmental context and then put them back into either the same context or a new context. Memory for the conditioned fear response was tested after 1 day and 28 days. Context sensitivity diminished at the longer, 28-day delay and the conditioned fear response was present in both familiar and novel environments. This suggests that the highly detailed elements of the memory trace (e.g., the details of the environmental context) are slowly stripped from the trace with the passage of time, in favour of a more gist-like representation that can be used to guide behaviour.

The Winocur et al. (2007) study provides support for the notion that passage of time can be particularly useful in gist abstraction and generalisation. This same pattern of results has been observed in human studies. With time, the regularities present in the representations are extracted whilst erroneous details are removed, leaving the basic gist of the event in LTM for later application to new scenarios (Sweegers & Talamini, 2014). This also leads to a decontextualisation of the memory traces over time whereby the target elements of an event are spared whilst the unnecessary supporting details are stripped away (Cox et al., 2014). Sweegers and Talamini (2014) had participants learn 144 face-location pairs, some of which were constrained by location rules (e.g., all faces appearing in a specific location belonged to the same category) that could be gradually extracted. Following the encoding session, participants either went directly on to the test, took a 2-hour nap, or engaged in 2-hours of quiet wakefulness. They were then tested on memory for arbitrary facial features, the location each face appeared in, and answered questions pertaining to whether they were explicitly aware of the location rules. Whilst the nap and quiet wakefulness group did not differ in terms of their ability to generalise and extract the location rules, they both showed increased ability to generalise when compared to the group that did not get a break between study and test, supporting the idea that generalisation takes time to emerge. Sweegers and Talmini (2014) suggest that the increase in generalisation comes at the cost of loss of details pertaining to the individual memory traces.

However, whilst increasing time between encoding and testing might benefit generalisation, it also increases the potential for forgetting. For example, the passage of time has been shown to significantly increase forgetting regardless of the number of stimuli presented or the number of times it was presented (Potter et al., 2002). Across their experiments, Potter and colleagues manipulated the length of time between study and test and observed a linear decrease in memory, indicating that forgetting significantly increases with time (Potter et al., 2002). If we want to decrease the chances of memories being forgotten, reactivation in the form of a second presentation of the stimuli or a retrieval of the memory trace might serve as forms of protection against forgetting. A more recent study presented participants with 300 images and tested their memory for some of the images immediately after encoding and after a 1-week delay. The extra retrieval led to better memory for those images after 1-week, suggesting that the extra encounter with the stimulus protected it from time-based forgetting (Mercer & Jones, 2019).

Mercer and Jones (2019) further supported this notion by running another experiment that had participants study some of the images twice, in an attempt to understand whether retrieval and/or exposure was inducing the protection from forgetting. They found that retrieval practice reduced the amount of forgetting across one week. Additionally, repeated exposure to the image during the study phase reduced forgetting throughout the week, but not to the same extent that the extra retrieval reduced forgetting (Mercer & Jones, 2019). Similarly, Nadel et al. (2007) examined the influence of repeated reactivations and retrievals across time on memory. They argue that repeatedly retrieving a target reactivates that memory trace and creates a new memory trace related to it. The extra memory traces therefore increase the amount of detail available for that event and reduces the amount of information forgotten as time passes. Further, they argue that reactivation may play a larger role in the fate of a memory trace than time (Nadel et al., 2007). More recently, Craig et al. (2021) presented participants with pictures of 60 item-location pairs. Half of the pairs were tested immediately and half were tested after a delay comprised of wakeful rest. They observed a protective effect of wakeful rest on the items that were not tested immediately following studying, arguing that the wakeful rest period protected those memory traces from forgetting (Craig et al., 2021). Regardless of whether sleep or wakeful rest follows encoding, no new information is being encoded which minimises the amount of interference that could contribute to items being forgotten. Alternatively, retrieval-induced forgetting might have contributed to their observed benefit of non-retrieved item. Retrieval-induced forgetting arises from a competition between related memories, leading to interference between the original memory trace and the memory traces that are reactivated with retrieval practice (Anderson & Hulbert, 2020).

The influence of time and forgetting on EM still warrants further investigation into the various ways it impacts our ability to retrieve episodic events. For example, Cox et al. (2014) compared memory for semantically congruent and semantically incongruent targets at two different time points. Forgetting was more severe for congruent targets at the longer time delay suggesting an interaction between both time and semantic expectations on EM processes. Further, Maxcey et al. (2018) postulated that recognition-induced forgetting arises when there is competition between semantically related objects. Correct recognition of one object, after it has had time to be consolidated into LTM, leads to an impairment in the ability to correctly recognise objects belonging to the same semantic category due to the amount of similarity between the inter-category objects (Maxcey et al., 2018). The final research goal therefore was to investigate what might be driving forgetting across time and its relation to memory performance in episodic memory. Specifically, how memory might be impacted by the interaction between factors such as context and semantic congruency across time. The aim was to identify ways to minimise forgetting and increase memory (e.g., via decontextualisation) as the length of time between study and test increases.

1.4 Conclusion

The literature reflects the complexity of the processes that influence episodic memory. Whilst one could argue that specific factors and processes exert an influence on how EMs are encoded or consolidated (e.g., congruency, context, time), it is too simplistic to examine them in solidarity and each factor should not be viewed as mutually exclusive. The overarching research question pertains to furthering our understanding of visual episodic memory, specifically the ways in which various factors discussed throughout this review interact with one another to influence how episodic memories are encoded, consolidated and retrieved.

Research Aims

- Examine the influence of semantic congruency and context-binding on episodic memory over time.
  - Do we see a long-term benefit of presenting targets in a consistently changing background and does that benefit persist for information that is semantically incongruent versus semantically congruent to the background scene?

- Examine whether verbal and visual episodic memory operate in the same manner.
  - Do the congruency and context-binding results observed in visual episodic memory apply to verbal episodic memory?

- Examine the neural markers of consolidation and context-binding in episodic memory.
  - Does theta activity during encoding, retrieval (or both) serve as a mechanism of context-binding?

2. Strength of Binding Determines Visual Long-term Memory for Objects

2.1 Introduction

Episodic memory is defined as detailed memory traces relating to specific events that are temporally grouped together (Baddeley, 1982). Findings suggest that large amounts of detailed information from the episode can be stored into visual episodic memory (Brady et al., 2008; Hollingworth & Henderson, 2002; Konkle et al., 2010; Shepard, 1967; Standing, 1973). What is more, the large amount of information that is stored in visual episodic memory seems to be highly detailed, going beyond just categorical information (Brady et al., 2008; Konkle et al., 2010). An aim of the present set of studies was to examine various factors that influence memory for images of objects in scenes.

One such factor is known as scene context. More broadly, context can refer to the physical or cognitive setting in which a stimulus or event is encountered (Tulving, 1983). When a visual episodic event is encoded, both the target and the contextual details of the event are encoded (Hollingworth, 2006, 2007). Context then supports accurate retrieval of an element within the experience (Hanczakowski et al., 2015; Hayes et al., 2007; Hollingworth, 2006; Jurewicz et al., 2016; Oliva & Torralba, 2007; Palmer, 1975). For example, Hollingworth (2006) showed participants a series of 40 images containing 7-11 to-be-remembered objects. Memory for specific objects in each scene was tested. Some of the objects were shown in their original scene whilst others were shown in isolation without any contextual information. They found that object recognition was highest when the objects were presented in a scene context (88.3% vs 79.8%). Further, better recognition for objects previously associated with a scene was observed when the test occurred after a delay in which another scene-object association was presented, suggesting that context facilitates robust visual memory. The elements of a visual scene are thought to be bound to the context in which they appear, forming a more comprehensive representation of the whole scene (Castelhano & Henderson, 2005; Criss & Howard, 2015; Hollingworth, 2006). Context aids memory regardless of whether or not the participant is aware that contextual associations are present (Castelhano & Henderson, 2005). More recently, Nakashima and Yokosawa (2011) replicated the findings of Hollingworth (2006), further supporting the idea that context aids memory retrieval, especially when large amounts of information are to be remembered. One could therefore argue that objects in visual scenes are bound to their context and that object-to-scene binding might serve as a memorisation strategy. This strategy could be particularly useful in situations in which a later memory test is expected since it facilitates the grouping of large amounts of associated information.

These findings complement another critical characteristic of context effects: the encoding specificity principle, which posits that the degree to which a retrieval cue matches the encoded material predicts the likelihood of successful retrieval (Tulving & Thomas, 1973). Due to this, we observe two opposing effects: 1) reinstating the same scene context results in improved recognition performance (i.e., context reinstatement) and 2) when the scene context changes between study and test, recognition memory decreases (i.e., context shift decrement). For example, reinstating the same scene context (e.g. during the study and test phases) in which target objects were presented in facilitated the recognition of target objects compared to when the accompanying scene was changed from study to test (Hanczakowski et al., 2015). Changing context between study and test causes significant interference with recognition memory in line with the encoding specificity principle due to the retrieval context (retrieval cue) not matching the encoding context.

If an encoding-retrieval cue match increases the likelihood of successful retrieval, then a decrease in accurate retrieval is expected when context cues change, reducing the benefit of context at retrieval. It has been argued that this decreases the likelihood of being able to accurately distinguish targets from distractors (Murnane et al., 1999). When a target is presented within a contextual scene (as opposed to being presented in isolation), the context is encoded regardless of whether it is relevant to the task or not (Brooks et al., 2010). The encoding of contextual information occurs automatically (Hayes et al., 2007) and is difficult to overcome. For example, Hayes et al. (2007) observed a context shift decrement effect despite participants being instructed to explicitly focus on the objects. They argue that the object-to-scene binding happens automatically, making it difficult to overcome the implicitly learned associations. This leads to a decrease in recognition (e.g., context shift decrement) when the anticipated associations change from study to test.

The decoupling of the tight relationships between an item and context has been postulated as one of the potentially necessary steps in memory consolidation. Referred to as the process of decontextualisation, it is defined as the gradual process of stripping detail from episodic memories so as to render them more abstract, gist-like memory traces. Decontextualisation is believed to work in a manner that weakens the associative links between the rich contextual details of the episodic memory over time (Cox et al., 2014; Jurewicz et al., 2016; Piolino et al., 2009; Winocur et al., 2007; Winocur & Moscovitch, 2011). It leads to loosening or breaking of the binding between the elements in an episodic event, leading to a more abstract representation of the event. It has been suggested that decontextualisation shows more robust effects in recognition than recall (Jurewicz et al., 2016; Piolino et al., 2009). Recognition tasks may be more sensitive to decontextualisation due to the fact that successful recognition does not rely on bringing to mind all of the details associated with a specific event compared to cued/free recall tasks which rely on bringing to mind the entire original memory trace. Whilst the extra details may also come to mind when engaging in recognition tasks, they are not required to successfully make old/new judgments (Mandler, 1980). It can be assumed that, if successful decontextualisation of a memory trace has taken place, then recognition memory should be better for those events due to the traces not being as reliant on details such as the context in which they were originally seen.

Recent research has attempted to manipulate the decontextualisation of targets to better understand how it influences memory and consolidation by attempting to force the decoupling of associative links. In these types of studies, targets are presented multiple times during a study session. The associated context is changed each time the target is encountered, weakening or removing any associations between the elements of the event before consolidation has taken place. With repeated exposure to a target, the original memory trace is reactivated, leaving it vulnerable to alteration. The way in which it is altered can take several forms such as strengthening the trace, changing the associations or even erasure of the memory (Hayes et al., 2007; Nadel et al., 2007, 2012; Piolino et al., 2009). In studies examining decontextualisation, the experimental paradigm exploits the idea that the memory trace can be altered by actively attempting to change the associative links to decrease the contextual dependence at retrieval as opposed to waiting for it to gradually happen over time (Evans & Wolfe, 2022).

If contextual details are incidentally encoded in order to facilitate the grouping of associated information, the question remains as to what underlies object-to-scene binding. Context may play a robust role in predicting what will be included in a scene due to expectations about what should be present in the scene (Oliva & Torralba, 2007). For example, one might expect a sofa to be present in a living room (e.g., semantically congruent) as opposed to a motorcycle in a bathroom (e.g., semantically incongruent). However, the effects of congruency on memory performance are debated, with some studies showing a memory benefit for congruent items and others showing a benefit for incongruent items (Cox et al., 2014; Gerver et al., 2020; Hollingworth, 2006; Pezdek et al., 1989; van Kesteren et al., 2020). For example, Friedman (1979) provided early evidence that semantic congruency aids memory for objects within a scene. They presented participants with a series of scenes, some of which contained objects typically encountered in that scene and objects that one would not normally encounter (i.e., semantically congruent with the background context and semantically incongruent with the background context). They observed an increase in response time during testing for objects that were congruent with their background scenes as compared to objects that were incongruent, suggesting that these objects required more attentional resources. They argue that the increase in reaction time was due to the fact that the congruent targets were highly likely to be found in those scenes, leading to increased processing time and more confusion about whether the target was actually present in that scene. Conversely, Davenport and Potter (2004) presented participants with semantically congruent and semantically incongruent object-scene pairs and found that recall accuracy was highest for objects that were congruent with their background contexts. This finding held true both when participants were told to ignore the background context and when they were asked to remember both the object and background. They propose that when an object is presented in a background context, the object and the scene are processed together. This associative binding occurs even when participants are told to ignore the background, aided by our expectations about a given scene context.

Assuming that objects and their background context are processed concurrently, unexpected objects (e.g., those that are semantically incongruent with their background context) may not be consolidated, leaving them prone to being forgotten due to a lack of association with their context (Cox et al., 2014). It is possible that congruent targets reactivate pre-existing associative networks, strengthening those memory traces and aiding later retrieval. Alternatively, it is possible that incongruent targets challenge the existing expectations, leading to the incongruent targets being retained in working memory for longer periods of time until new associations are formed (Pezdek et al., 1989). Greve et al. (2019) argue that congruency operates on a ‘U-shaped function’ in that highly congruent and highly incongruent events will both be better remembered than mildly incongruent events. Highly congruent targets are thought to activate pre-existing schemas which leads to more efficient encoding. On the other hand, highly incongruent targets violate schema expectancy, leading to more efficient encoding due to a need for updating schemas in the presence of a continually changing environment (Greve et al., 2019).

This therefore warrants further testing of the way in which elements of visual scenes are bound together to support accurate recognition of only one of the elements at test. Importantly, Evans and Wolfe (2022) sought to examine under which conditions context may play a negative role in memory, contrary to the vast amount of literature demonstrating positive benefits of context. Across a series of 12 experiments, they tested memory for objects in isolation, objects in arrays and objects in scenes (for experimental conditions and results see Figure 2.1 below). The general experimental design employed across their 12 experiments involved presenting participants with a series of 100 target objects for 3 seconds each during the study phase, followed by an immediate old/new recognition test for all of the target objects. Participants were shown the target objects twice during encoding only in Experiment 10, otherwise, target objects were shown only once in each scene during encoding.

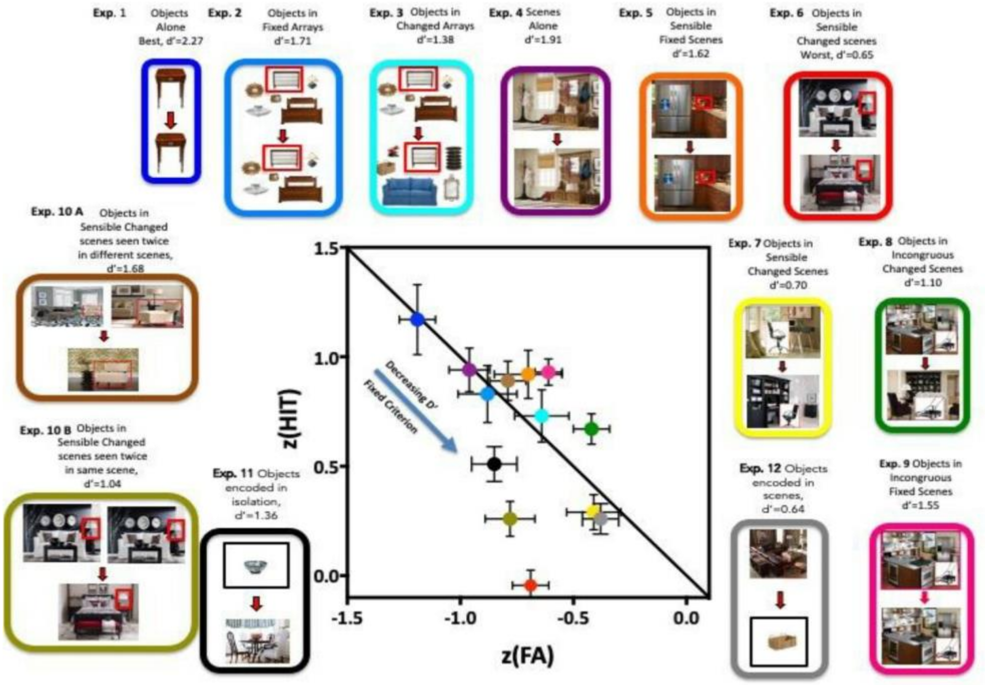

Figure 2.1 Experimental conditions and results for each of the 12 experiments from Evans and Wolfe (2022). Better performance is indicated by points closer to the upper left corner of the figure and the diagonal line indicates a neutral decision criterion.

Evans and Wolfe (2022) base their argument of negative context effects on the finding that performance across the 12 experiments never reached the same level as found when objects were both studied and tested in isolation (see Experiment 1 in Figure 2.1). However, several aspects of episodic memory were not taken into consideration. For example, they posit that the context, even though irrelevant to the task, crowded the representation of the object. This was based on a review of literature examining crowding effects by Levi (2008), but it seems that this is argument by analogy rather than a case for assuming that visual crowding is the actual cause of the memory decrement. Evans and Wolfe (2022) argue "that tight, automatic associations between target items and sensible backgrounds are crowding or hiding the mnemonic representation of those [congruent targets appearing in changing scenes] targets" (p. 238). The authors make this argument in passing, providing no explanation as to what processes might be driving this effect. Further, this argument is based on a specific instance of when a semantically congruent target object is studied in one scene-context and then tested in another. Overall, therefore many questions remain regarding how object-scene congruence influences memory of particular target objects. For instance, it remains unknown where in the cognitive system the disruption to memory occurs (e.g., does it occur at encoding, at retrieval or both?).

Another consideration that was not taken into account is how semantic congruency is operating in conjunction with context effects. Evans and Wolfe (2022) conducted two experiments to test whether context influences memory for targets that were semantically incongruent with their background context. In Experiment 8, participants were shown 100 images and asked to memorise a single object, appearing in a red bounding box, that was semantically incongruent with the scene. Memory was tested for each of the objects, however the objects appeared in a new scene at test. They found that semantically incongruent targets were better remembered than semantically congruent targets, aligning with the wider field of studies on congruency effects. However, their comparisons were across two separate experiments (Experiments 7 and 8; Experiment 7 was identical to Experiment 8 except that, in Experiment 7, the targets were semantically congruent with the scene context). When the scene changed from study to test memory was better for incongruent than congruent targets. In Experiment 9, however, they tested memory for incongruent targets that appeared in a reinstated context at test (again, using 100 scenes/objects). When the scene context was reinstated, memory for the incongruent targets was better than when the incongruent targets changed scenes between study and test (Experiment 8 versus 9). But now the memory benefit for incongruent targets was abolished. Comparing across Experiment 9 with the congruent equivalent experiment (i.e., Experiment 5), performance was found to be similar for congruent and incongruent targets. In sum the evidence regarding the influence of object/scene congruence is mixed: there is no congruency effect when the context is reinstated (Experiments 5 vs 9), yet there is a benefit for incongruent targets when the context changes between study and test (Experiments 7 vs 8). It is therefore unclear, based on the experiments and findings from Evans and Wolfe (2022) how semantic congruency is operating in conjunction with context effects in visual episodic memory.

Evans and Wolfe (2022) further attempted to strengthen the memory traces and ‘decontextualise’ them by presenting the targets multiple times during the study phase before testing memory for the objects. However, this was only manipulated for the semantically congruent targets. In Experiments 10A and 10B, congruent targets were each presented twice during the encoding phase. In Experiment 10B, the scene was held constant at encoding and changed at test. In Experiment 10A, the scene changed between the two presentations at encoding, and there was a third, new scene at test. Memory for the targets was incredibly poor when the target appeared in the same scene twice and in a new scene at test (Experiment 10B). When comparing memory for targets between Experiments 10A and 10B, memory was better when the target was encountered in different scenes of the same type (e.g., a lamp in two different bedrooms) than when the same object-scene image was repeated twice at encoding. This suggests that the memory for targets encountered in different scenes of the same type had become more abstract or ‘gist’ like. Indeed there is evidence to suggest that re-encountering a target object (via multiple presentations) leads to a ‘re-encoding’ and ‘reconsolidation’ of the target (Zöllner et al., 2023). It is important to note that re-encoding a target in changing scenes and re-encoding/reconsolidation are two different processes, both of which might be contributing to gist representations within the design of the present experiments. Evidence suggests that retrieval of a past episode leads to activation of related memory traces, reactivating the related traces, facilitating the organisation of new information, and in some cases leading to reconsolidation (Zöllner et al., 2023).

It is therefore possible that when a target object is presented in a new scene, the original scene-context is reactivated, strengthening the representation of the target object by updating the memory trace to include both scenes. This would lead to a more ‘gist’ like representation of the target, less dependent on a specific scene context. Encoding variability posits that repeating a target in more than one context promotes distributed representations of the to-be-remembered item, facilitating retrieval when the target appears in a similar context (Imai & Richman, 1991). Alternatively, if a target object is encountered in changing scenes, then it is possible that the representation stored is more schematic in nature (i.e., a chair in a living room rather than a specific chair in that room). Previously established schemas that provide expectancies about what should be present in a scene-context are believed to aid encoding by retaining the simplest information needed to convey the schema of a complex image (Pezdek et al., 1988). The dynamic memory model posits that episodic memories are stored based on pre-existing knowledge structures (i.e., schemas) which are activated when processing an original event (Lampinen et al., 2001). This can lead to schema-relevant information (i.e., congruent information) being better remembered at retrieval due to a better integration with what is already present in the existing schemas

(Brewer & Treyens, 1981). Therefore, when the target objects are presented in changing scenes during encoding, it is plausible to assume that what is being stored is the more schematised representation of the target, creating a more gist-like representation, that should promote memory for the episode if it is congruent with expectations of that scene.

Further evidence has shown that context reinstatement facilitates memory for congruent items to a greater extent than incongruent items and that changing contexts hinders memory for congruent targets to a greater extent than incongruent targets due to congruent items being better integrated into memory traces (Bein et al., 2015). Additionally, when a target object appears more than once at encoding and the context changes, the distinctiveness of the memory traces for that target object become blurred (Laurent et al., 2020). It has been postulated that the distinctiveness of the memory traces resulting from multiple presentations should be higher for congruent targets compared to incongruent targets. It remains unclear what the effect of repeating object/scene images has on memory for incongruent cases. Indeed, it remains possible that the pattern Evans and Wolfe (2022) observed in Experiments 10A and 10B might not be the same if using semantically incongruent targets. One could expect performance to be similar, in line with their comparison between Experiments 5 vs 9. Therefore, not only does semantic congruency warrant further examination, but it is also necessary to better understand how it interacts with context effects after repeated presentations.

Finally, the time needed for consolidation was not considered. In each of their experiments, the old/new test followed immediately after the study phase. To truly examine the role context plays in visual episodic memory, the time needed for consolidation must be taken into account. For example, the standard/systems-level consolidation theory posits that consolidation occurs during ‘offline’ periods (e.g., sleep) that facilitate the transfer of newly encoded memories from the hippocampus to the neocortex (Dudai, 2004; Reiner et al., 2014; Yonelinas et al., 2019). Whilst a period of sleep is not always necessary for consolidation, there is time needed for consolidation to take place. Studies have shown that this can be as short as a 15 min period of quiet wake so long as there is a reduction in attentional demands and new learning (Wamsley, 2019). Episodic memory has been shown to be dependent upon the length of time between study/test, with performance dropping with long delays (Zöllner et al., 2023). The question remains as to whether the repeated presentation combined with a longer test delay can preserve the memory traces.

The initial study described here sought to better understand the role that context and semantic congruency plays in visual recognition memory by extending the work of Evans and Wolfe (2022). The aim of the study was to examine in more detail the influence of both object-scene congruence and context on the recognition memory of individual objects. The study comprised three main context conditions: No Change, Contextualised and Decontextualised. The No Change condition extended Experiment 5 from Evans and Wolfe (2022). In this condition, participants were shown a target object twice during encoding. Each time the target was presented, the scene context was held constant. The target object was then presented in the same context during test. Our Contextualised condition was an extension of their Experiment 10B. Like the No Change targets, objects appeared in the same scene context both times they were encountered during encoding, but they were placed in a new context at test. Finally, the Decontextualised condition was a replication of Experiment 10A. In this condition, the targets were always encountered in a new scene context. The current study also sought to extend the methods from Evans and Wolfe (2022) by examining how semantic congruency operates within each of the three context conditions. Therefore, for each context condition, participants were presented with an equal number of semantically congruent and semantically incongruent targets. This was done to extend the limited findings from their Experiments 8 and 9 and to clarify how congruency works within the framework of context reinstatement and changing contexts.

2.2 Methods

2.2.1 Participants

Data were collected from 56 healthy adults aged 18-30 years (43 Female). The sample size was in excess of minimum of 12 as tested in the experiments reported by Evans and Wolfe (2022). Six participants were excluded due to technical difficulties, six were excluded due to missing a testing session, two were excluded for performing at chance after the immediate test session, one was excluded for completing the delayed test outside the 24-hr window, and one was excluded for failing the attention checks. This left us with a final sample of 40 participants. All participants were native English speakers, had normal or corrected-to-normal vision, had no special visual characteristics (e.g., colour blindness or astigmatism), and had no history of drug or alcohol abuse. They had no history of hearing problems, dyslexia, or neurological, psychiatric or sleep disorders, and they did not have a job that involved alternating shift patterns (i.e., day and night shifts). All participants provided informed consent and filled in demographic questionnaires before the study. The current study was approved by the Research Ethics Committees of the Department of Psychology, University of York.

2.2.2 Materials

2.2.2.1 Images

Stimuli consisted of 822 (targets and foils) unique, coloured images gathered via Google Image searches. Target objects were embedded in background contexts and identified using a red bounding box. Each target object appeared within a real-world indoor scene (e.g., living rooms, kitchens, bedrooms). The 216 target objects were either semantically congruent with the background scene (e.g., a kettle in a kitchen; 50% of targets) or semantically incongruent with the background scene (e.g., a kayak in a bedroom; 50% of targets). Figure 2.2 provides an example of the congruent and incongruent stimuli and Table 2.1 outlines the categories included in our stimuli set for each type of target. To make the congruent images, images were collected from image searches where the target object was already present in the image, and which had the target object in multiple indoor scenes. Incongruent targets were objects that were edited to remove their backgrounds using Preview for macOS Catalina 10.16.6 and then superimposed onto an indoor scene. Each image was resized to 800 x 700 pixels, though due to the experiment being conducted online it is difficult to specify the degrees of visual angle or the viewing distance for each participant.

Table 2.1 Exemplar categories for congruent and incongruent target objects.

Congruent Object Categories	Incongruent Object Categories
Appliances	Anchors
Chairs	ATVs
Sofas	Bicycles
Beds	Cars
Cabinets	Kayaks
Desks	Lawnmowers
Lamps	Machinery
Chandeliers	Motorcycles
Tables	Scooters
Ornaments	Tricycles

Figure 2.2 Example of congruent and incongruent targets within indoor scenes.

2.2.2.2 Experimental conditions in the experiment

In general terms the experiment was designed around an initial study phase followed by three later testing phases. Each of the target objects, half of which were Congruent and half Incongruent, appeared twice in the study phase. The order of the appearance of the targets was randomised and was different for each participant. In each testing phase the participant engaged in series of Old/New trials in which on each trial a single image was presented and the participant had to decide whether the designated target had been presented during the study phase. The targets were randomly assigned to one of the three different testing phases intermixed with foil objects embedded as targets in background scenes. An equal number of target objects were assigned to three context conditions that were intermixed and their order of appearance was randomised; No Change, Contextualised and Decontextualised conditions (see Figure 2.3). In the No Change condition, the target appeared in the same background scene both times it was presented during the study phase. The same image was then presented at test. In the Contextualised condition, the target appeared in the same background scene both times during the study phase but was placed in a new background at test. Finally, the Decontextualised targets appeared in a different background context each time it was presented during the study phase and in a third, new background during at test.

Each of the three test conditions occurred immediately following the study phase (at Immediate test phase), again 24 hours later (at the Delayed test phase) and one week later (at the Follow up test phase). Targets were randomly assigned to one of the three testing phases: Immediate, Delayed and Follow-up. Each testing phase consisted of 72 targets from the study phase plus 72 foils for a total of 144 trials per testing phase. Hence, all of the target objects from the study phase were tested during one of the three testing phases, with memory for each target being tested only once during the study. The choice of which target object was tested at each testing phase was randomly assigned for each participant. However, in each testing phase there was an equal number of congruent and incongruent targets as well as an equal number of targets from the three context conditions (12 per context condition that were Congruent and 12 per context condition that were Incongruent).

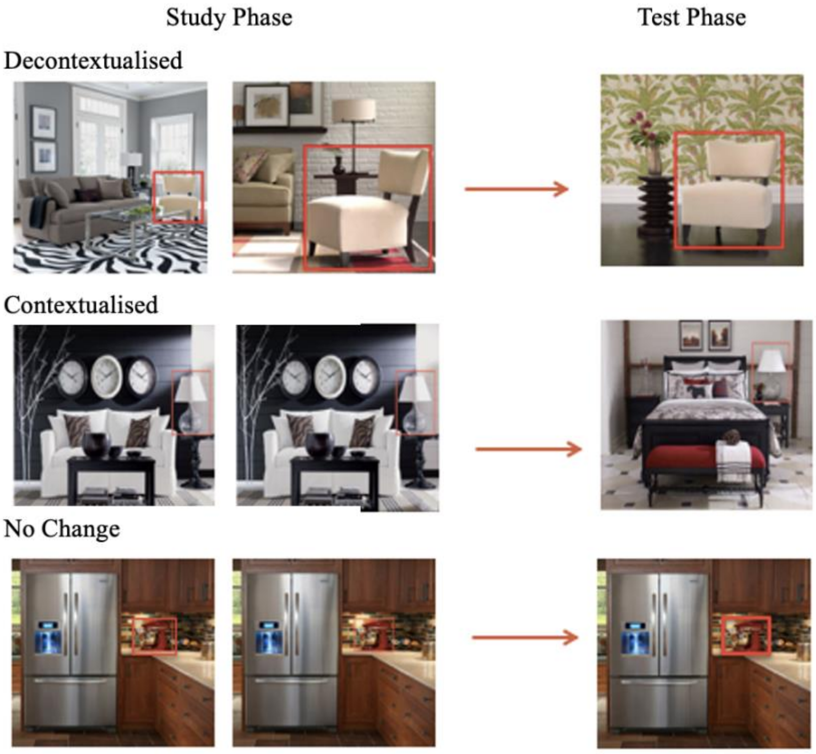

Figure 2.3 Example images for each of the three conditions: Decontextualised, Contextualised and No Change.

2.2.2.3 Experimental Design

The experiment was conducted using a within-subjects design. The three main factors examined were Context (No Change, Contextualised, Decontextualised), Congruency (Congruent, Incongruent) and Time (Immediate, 24-h Delayed, 1-wk Follow Up). Within each of the three context conditions, there were an equal number of congruent and incongruent target objects. Each unique target was tested once during the experiment as was randomly assigned to one of the three testing sessions.

2.2.3 Procedure

The experiment was conducted online. The experiment was coded in Visual Studio Code using JavaScript, HTML, C# and CSS. It was hosted on Microsoft Azure and participants were sent a link to the experimental web page when they signed up to participate. The experimental code and stimuli set is available on GitHub (https://github.com/evmadden/SDCexp1).

Participants were instructed to use Google Chrome for the study. The website would not allow participants to enter the study if they were using a mobile phone or tablet (determined by detecting the screen size of the browser they were using). They were given instructions on how to put their browser into full-screen mode, with the address bar hidden. If the browser was removed from full screen at any point during testing, the experiment stopped, and they were asked to contact the researcher. This was done to ensure that images were sized correctly (800 x 700 pixels) and were fully visible in the centre of the screen.

Participants were asked to complete an encoding phase followed by an immediate, baseline memory test over Zoom with the experimenter. They returned to the website 24-hours later, at the same time they completed the previous session, to complete the first delayed memory test. Finally, they were asked to return to the website 1-week after the study phase/immediate test, at the same time they completed the first two sessions, to complete the final, follow-up memory test (see Figure 2.4). Reminder emails were sent to the participants ahead of each session and the website would not allow them to start the next session until the appropriate amount of time had passed.

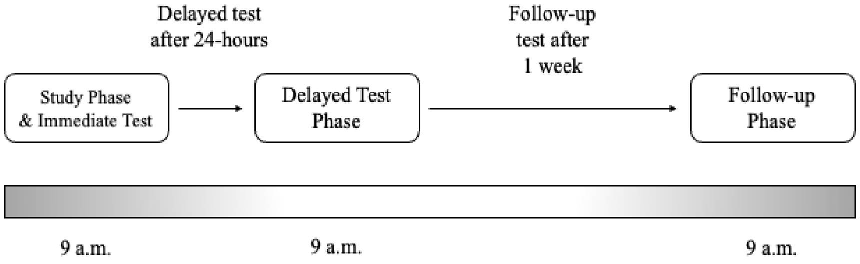

Figure 2.4 Diagram showing the experimental timeline for each phase of the experiment.

Participants first provided informed consent and answered basic demographic questions. They were then asked to complete the Stanford Sleepiness Scale (Shahid et al., 2011). During the first session, participants completed both the study phase and the initial test phase, whereas they only completed a test session at the 24-hour and 1-week sessions.

2.2.3.1 Study Phase

During the study phase, images were presented in the centre of the screen for 3 seconds each followed by a fixation cross for 1.5 seconds. Participants were presented with a single sequence of 432 coloured images (216 unique targets, each presented twice). Participants were instructed to only pay attention to the objects within the box and memorise it, ignoring the scene. As an attention check participants were asked to make a manual response when a number appeared on the screen instead of an image. Any time a number appeared on the screen the participant was asked to press the 'spacebar'. If the participant did not respond within 5 seconds, the next image was presented. Throughout the study phase, there was a 10% chance that the participant was shown a number instead of a picture (see Figure 2.5). At the end of the study phase, if the participant had failed an attention check they were prompted to give a reason as to why. Each failed attention check was logged alongside any instances of the web browser exiting full screen mode.

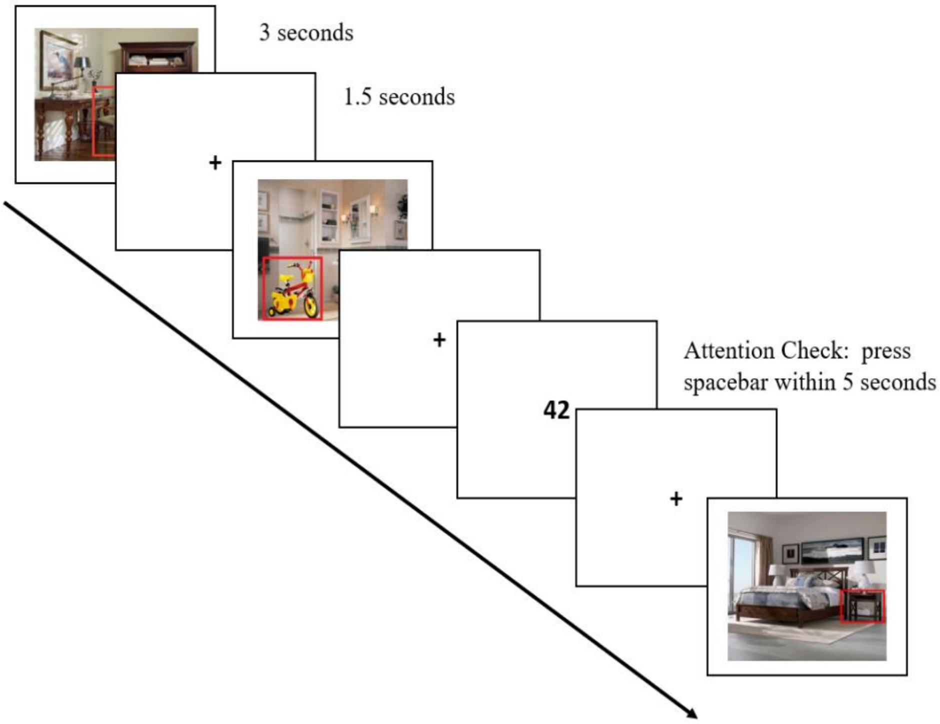

Figure 2.5 Example of study phase sequence. Image was centred on the screen and presented for 3 seconds, followed by a fixation cross for 1.5 seconds. There was a 10% chance that a number, as opposed to a picture, would appear on the screen (attention check). If a number appeared on the screen, participants had 5 seconds to press the spacebar and continue to the next image.

2.2.3.2 Testing Phases

For each test trial, participants were shown an image and asked to judge whether it was 'old' or 'new' by pressing the left and right arrow keys respectively. Response time for the judgement was recorded in milliseconds. Once the old/new judgement was made, they were asked to rate how confident they were in their answer on a 4-point scale: 1 - not very confident, 2 - somewhat confident, 3 - confident, 4 - very confident. Finally, participants were given feedback on whether their answer was correct or incorrect (See Figure 2.6), with feedback remaining on the screen for 1.5 seconds before progressing to the next test trial. Each test session was self-paced. After the final testing phase, participants were asked to complete the Pittsburgh Sleep Quality Index (Buysse et al., 1989) and the Epworth Sleepiness Scale (Johns, 1991). Finally, participants were debriefed and instructed to contact the researchers if they had any further questions.

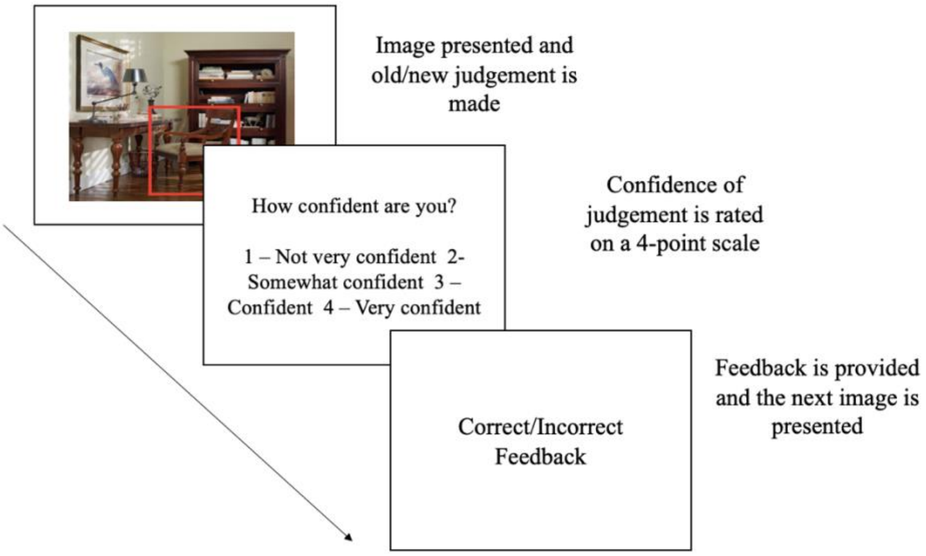

Figure 2.6 Example of test trial procedure. Participants were shown an image, which remained on the screen until a response was made. Participants then rated how confident they were in their answers and feedback on their answer was provided.

2.3 Results

2.3.1 Hits and False Alarms

The first step to understanding the current findings was to check to see whether the results at the Immediate test replicated previous work by Evans and Wolfe (2022) that the current study was based on (see Figure 2.7) by plotting the z-transformed hit rate against the z-transformed false alarm rate (for the comparable figure see Figure 10 in Evans & Wolfe, 2022). A similar trend in the pattern of results was observed despite the absolute d-prime (d’) values differing (see Table 2.2). When the context is reinstated, performance is best (regardless of whether the target is semantically congruent or semantically incongruent with the scene context) and changing the context leads to a steep decline in performance consistent with the classic context reinstatement and context shift decrement effects. However, if the target is consistently encountered in a new scene context (Decontextualised condition), there was a slight rebound in performance relative to the Contextualised condition, for the Congruent targets but a further decrease in performance for the Incongruent targets. This pattern of results led to the examination whether the differences observed in the data were significant (discussed below).

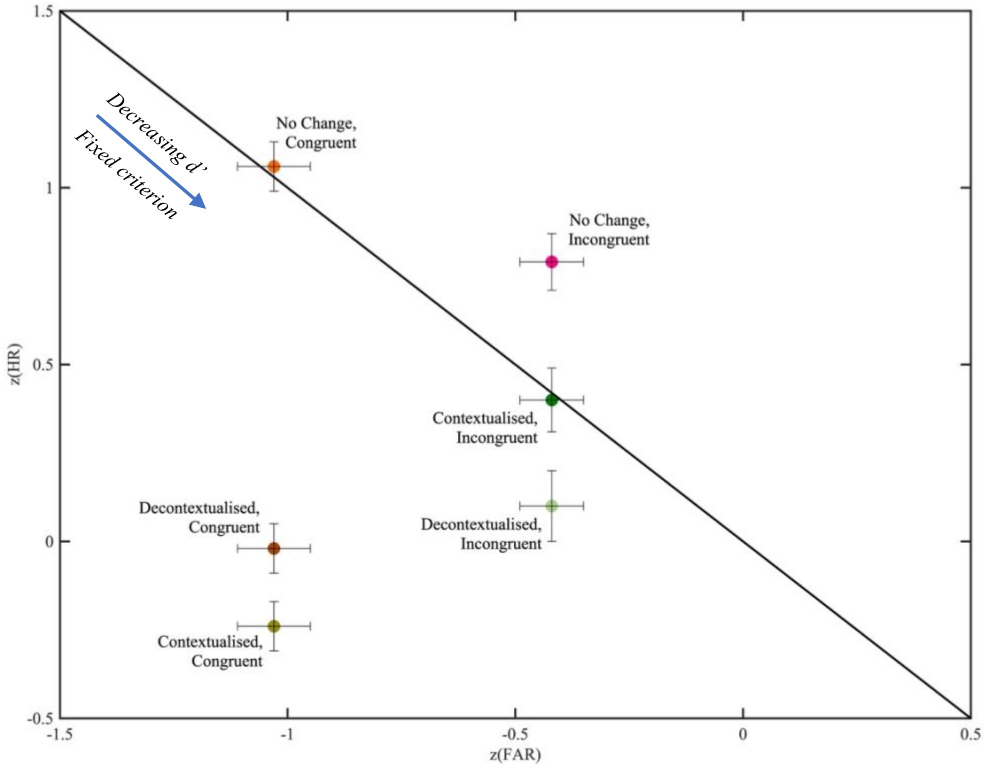

Figure 2.7 Memory performance at the Immediate test represented by z-transformed Hit Rate plotted against z-transformed False Alarm Rate. Better performance is indicated by proximity to the upper left corner and the diagonal black line indicates a neutral criterion. Each data point represents the average for one condition in the experiment. Error bars represent ±1 *SEM*.

Table 2.2 Average d’ scores for each condition in our experiment and from the corresponding experiment in Evans and Wolfe (2022).

Condition	Current Study (d’)	Experiment	Evans & Wolfe (d’)
Congruent		Congruent	
No Change	2.09	Experiment 5	1.62
Contextualised	.79	Experiment 10b	1.04
Decontextualised	1.01	Experiment 10a	1.68
Incongruent		Incongruent	
No Change	1.21	Experiment 9	1.55
Contextualised	.82	Experiment 8	1.10
Decontextualised	.52		

Note. Experiments 5, 9 and 8 from Evans and Wolfe (2022) only presented targets once during encoding as opposed to the current study which presented targets twice during encoding. There was no comparable Evans and Wolfe (2022) experiment for the Incongruent Decontextualised condition in the present study.

2.3.2 D-Prime and Criterion

The foils at test were always embedded in new, unique background scene contexts and different foils were used on the congruent and incongruent test trials respectively. As a consequence, different false alarm rates occurred across the congruent and incongruent conditions, meaning statistical comparisons between the three context conditions could not be assessed. However, it allowed d’ scores to be calculated for the congruent and incongruent targets separately within each of the context conditions. These allowed us to conduct three separate 2 x 3 repeated-measures ANOVAs with the factors *Congruency* (Congruent, Incongruent) and *Time* (Immediate, 24-hr, 1-wk) for each of the three context conditions using d’, c, and separately averaged confidence. Negative c values indicate a more liberal criterion (i.e., increased likelihood of responding ‘old’ to a target) whilst positive values indicate a more conservative criterion (i.e., decreased likelihood of responding ‘old’ to a target).

2.3.2.1 No Change Condition

We first examined the effects of congruency and time on targets that were consistently shown in the same scene context (i.e., No Change condition; context reinstatement) via d'. We observed main effects of *Time* (*F*(2, 78) = 52.86, $p < .001$, $\eta_p^2 = .58$) and *Congruency* (*F*(1, 39) = 61.72, $p < .001$, $\eta_p^2 = .61$). Recognition linearly decreased across the week (*F*(1, 39) = 138.30, $p < .001$, $\eta_p^2 = .78$) and Congruent targets were significantly better remembered than Incongruent targets (See Figure 2.8). Additionally, we observed a significant *Congruency X Time* interaction (*F*(2, 78) = 8.07, $p = .001$, $\eta_p^2 = .17$). Simple main effects showed that at each test session, participants were significantly better at recognising congruent targets compared to incongruent targets (all time points, $p \leq .001$).

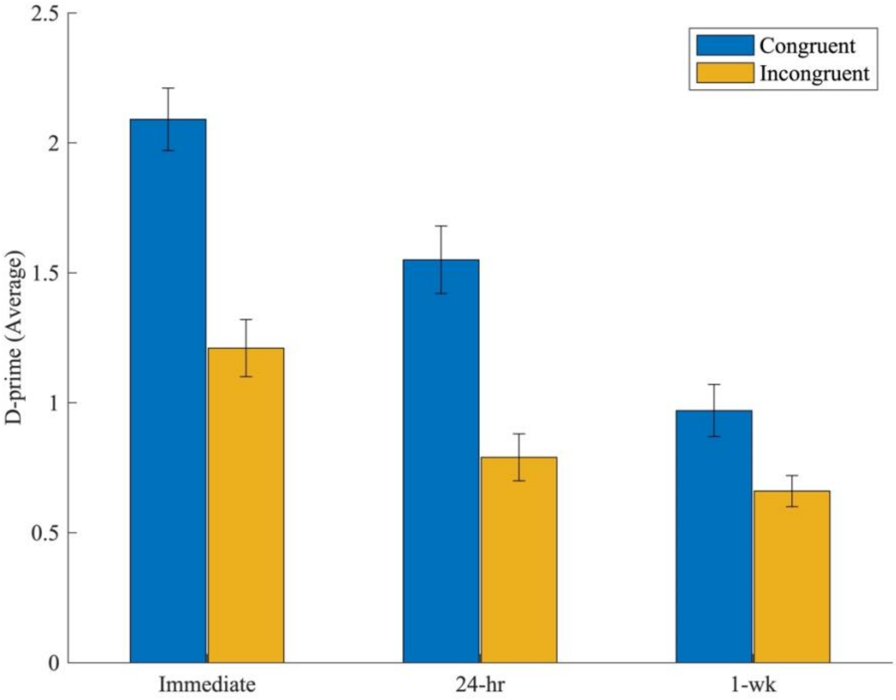

Figure 2.8 Average d' scores for target memory in the No Change condition at each test session (Immediate, 24-hour delay, 1-week delay) based on congruency (Congruent and Incongruent). Error bars represent ±1 SEM.

Furthermore, when examining the response bias in the No Change condition, there were main effects of *Congruency* (*F*(1, 39) = 3.13, $p = .049$, $\eta_p^2 = .07$) and *Time* (*F*(2, 78) = 49.45, $p < .001$, $\eta_p^2 = .56$). Participants tended to adopt a more conservative criterion threshold for the Congruent targets (*M*

= .14, *SE* = .04) when compared with the Incongruent targets (*M* = -.20, *SE* = .05). Participants were liberally responding at the Immediate (*M* = -.10, *SE* = .04) and 24-h delayed (*M* = -.04, *SE* = .05) tests, but applied a more conservative criterion at the 1-wk follow-up test (*M* = .06, *SE* = .07). Finally, there was a significant *Congruency X Time* interaction (*F*(2, 78) = 4.87, *p* = .010, η_p^2 = .11, see Figure 2.9). Simple main effects of interaction showed that the effect of congruency was present at all time points (Immediate: *p* = .008; 24-h and 1-wk: *p* < .001) and as Figure 2.9 shows this tended to increase as the lag between study and test increased.

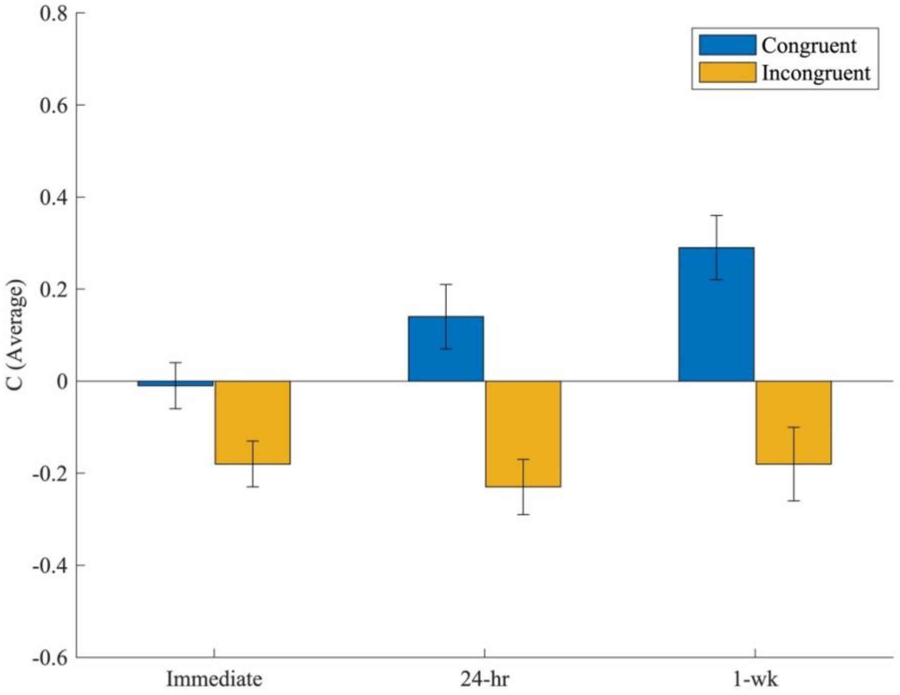

Figure 2.9 Average criterion for old/new decision in the No Change condition at each test session (Immediate, 24-hour delay, 1-week delay) based on congruency (Congruent and Incongruent). Error bars represent ±1 SEM.

2.3.2.2 Contextualised Condition

We next examined the effects of congruency and time on targets that were shown in one context during study and in a new context during test (i.e., Contextualised condition; context shift decrement) based on d'. In this condition, we observed a main effect of *Time* (*F*(2, 78) = 10.77, *p* < .001, η_p^2 = .22) but not *Congruency* (*F*(1, 39) = 1.28, *p* = .266, η_p^2 = .03). Like with the No Change targets, recognition linearly decreased across the week (*F*(1, 39) = 20.78, *p* < .001, η_p^2 = .35). However, there was no difference between Congruent and Incongruent targets. Interestingly, we observed a significant

Congruency X Time interaction ($F(2, 78) = 3.26$, $p = .044$, $\eta_p^2 = .08$). Whilst there is no difference between congruent and incongruent targets, as a whole, simple main effects of interaction showed that there was a slight difference in performance between the two target types at the delayed test (see Figure 2.10).

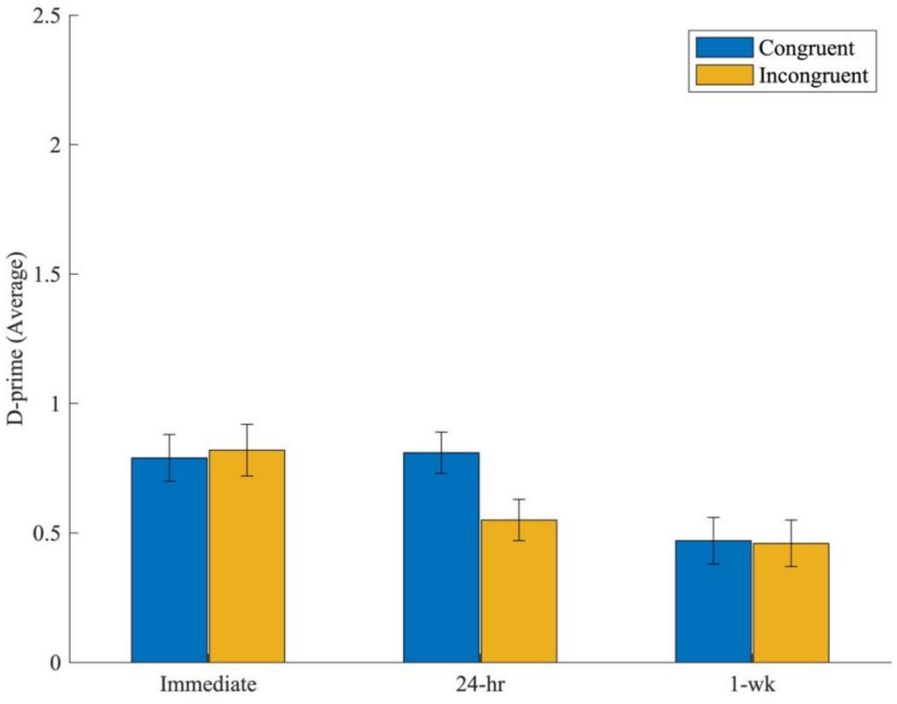

Figure 2.10 Average d' scores for target memory in the Contextualised condition at each test session (Immediate, 24-hour delay, 1-week delay) based on congruency (Congruent and Incongruent). Error bars represent ±1 SEM.

Examination of the criterion threshold for targets whose scene context changed between study and test revealed there was a main effect of *Congruency* ($F(1, 39) = 86.80$, $p < .001$, $\eta_p^2 = .69$) but no main effect of *Time* ($F(2, 78) = 1.57$, $p = .215$, $\eta_p^2 = .04$). Much like No Change targets, participants adopted a conservative criterion threshold for the Congruent targets ($M = .56$, $SD = .05$) whilst they applied a liberal criterion for the Incongruent targets ($M = -.06$, $SD = .05$). Furthermore, the *Congruency X Time* interaction ($F(2, 78) < .01$, $p = .996$, $\eta_p^2 < .01$) was found to not be significant (see Figure 2.11) . When there is a change in context between study and test, participants respond more conservatively to congruent targets and respond more liberally to incongruent targets regardless of when the test occurs. This is likely to reflect the increased difficulty in identifying the incongruent targets when the retrieval cue is no longer helpful.

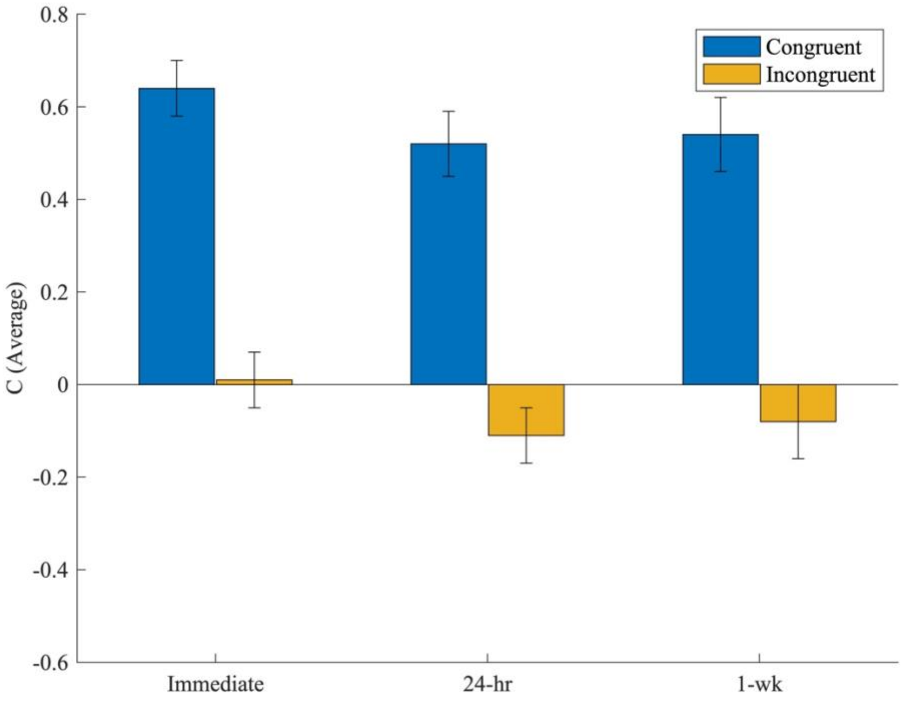

Figure 2.11 Average criterion for old/new decision in the Contextualised condition at each test session (Immediate, 24-hour delay, 1-week delay) based on congruency (Congruent and Incongruent). Error bars represent ±1 SEM.

2.3.2.3 Decontextualised Condition

Finally, we examined the effects of congruency and time on targets that were consistently shown in a new scene context (i.e., Decontextualised condition), based on d' scores. In this condition, we observed main effects of *Time* ($F(2, 78) = 8.92$, $p < .001$, $\eta_p^2 = .18$) and *Congruency* ($F(1, 39) = 39.46$, $p < .001$, $\eta_p^2 = .50$), in line with the findings from the No Change condition. Again, recognition linearly decreased across the week ($F(1, 39) = 18.73$, $p < .001$, $\eta_p^2 = .32$) and Congruent targets were better remembered than Incongruent targets. However, there was not a significant *Congruency X Time* interaction ($F(2, 78) = 2.71$, $p = .073$, $\eta_p^2 = .07$; see Figure 2.12). At each test session, memory for Congruent targets is better than for Incongruent targets, and memory performance in both conditions decreases across the week.

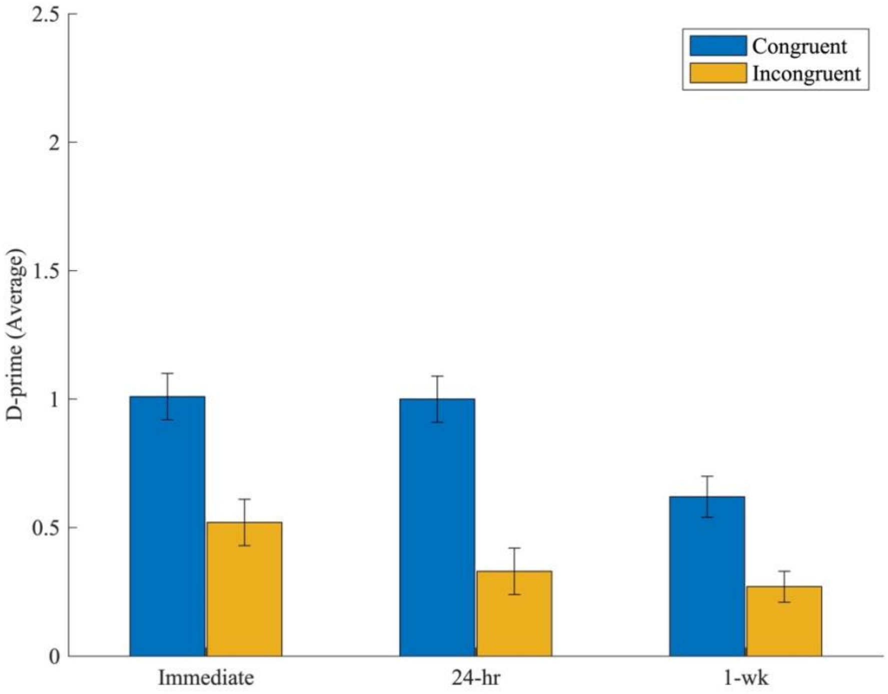

Figure 2.12 Average d’ scores for target memory in the Decontextualised condition at each test session (Immediate, 24-hour delay, 1-week delay) based on congruency (Congruent and Incongruent). Error bars represent ±1 SEM.

Furthermore, the results of the criterion analysis for the Decontextualised targets, much like the Contextualised targets, demonstrated there was a main effect of *Congruency* ($F(1, 39) = 49.90$, $p < .001$, $\eta_p^2 = .56$) but no main effect of *Time* ($F(2, 78) = 2.42$, $p = .096$, $\eta_p^2 = .06$). When the targets were presented in continually changing scenes, participants adopted a conservative criterion threshold. However, the criterion for the Congruent targets ($M = .47$, $SD = .05$) was more conservative than for the Incongruent targets ($M = .06$, $SD = .05$). Again, the *Congruency X Time* interaction ($F(2, 78) = .41$, $p = .663$, $\eta_p^2 = .01$) was found to not be significant (see Figure 2.13).

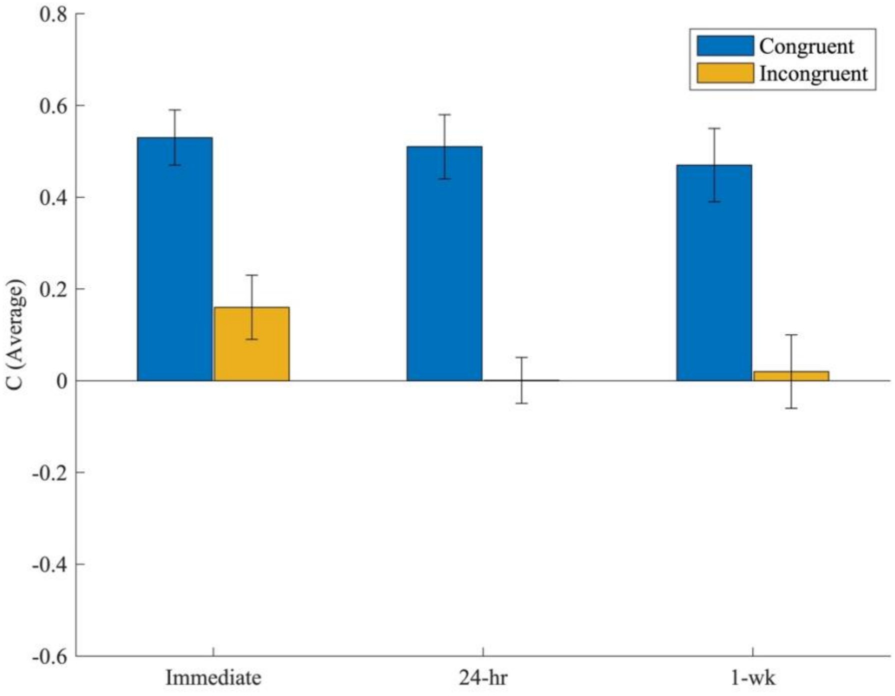

Figure 2.13 Average criterion for old/new decision in the Decontextualised condition at each test session (Immediate, 24-hour delay, 1-week delay) based on congruency (Congruent and Incongruent). Error bars represent ±1 SEM.

2.3.3 Confidence Ratings

Confidence ratings were taken so as to see whether these might shed further light on memory for congruent vs incongruent targets. Participants, at test, rated how confident they were in their answers on a scale ranging from 1 – not at all confident to 4 – very confident. Reinitz et al. (2011, 2012) argue that the strength of a memory trace determines how accurate participants will be in their old/new responses whilst certainty, as well as item strength, contributes to how confident a participant is in their response. This view aligns with dual process theory which posits that recollection and familiarity contribute to accuracy in recognition decisions (Deason et al., 2012). Within this framework, recognition refers to specifically recalling an event and bringing all of the details to mind whilst familiarity relates to a more general feeling of previously encountering the event without bringing to mind specific details (Deason et al., 2012). Reinitz et al. (2011, 2012) argue that the strength of the memory trace is therefore determined by successful encoding/retrieval whilst certainty arises as a result of successful retrieval of an encoded item in combination with whether or not the participant fully recalled the specific item, in line with dual-process theories. This theoretically results in higher confidence for specific features of an image (e.g., remembering a specific chair) compared to general familiarity with an image and higher accuracy for familiarity with an image when compared to specific features (e.g., remembering seeing a chair in a sitting room; Reinitz et al., 2011, 2012).

In their studies, feature-based stimuli included pictures with specific objects that aided retrieval of the whole image whilst familiarity-based stimuli consisted of images that did not have a distinct object, leading to retrieval being based more on being familiar with the image rather than specific retrieval of an object. Which condition the image belonged to in their study was determined by asking participants to rate whether their recognition response was based on a specific feature within the image or based on general familiarity with the image as a whole (Reinitz et al., 2011). Within this theoretical framework, and by analogy, Congruent targets in the present study were assumed to be akin to the familiarity-based targets (given they were in expected scenes), and Incongruent targets were considered feature-based targets (since they were not related to the scenes). Accuracy and confidence raitings were corrected for FAs using the following formuals: Corrected Accuracy = (H – FA)/(1-FA) and Corrected Confidence = H – FA, where H corresponds to the average accuracy/confidence of hits and FA corresponds to the average accuracy/confidence of the false alarms (Reinitz et al., 2011, 2012). These measures were calculated separately for Congruent and Incongruent targets at each time point and for each context condition.

The results from the confidence raiting were, in general, uninformative when considering the results alongside those of d’ and c, despite the promising theoretical framework provided by Reinitz et al. (2011, 2012). The only interesting finding in terms of confidence was that participants were more confident when making FAs compared to Hits, evidenced by the negative confidence scores when correcting for FAs (see Figure 2.14).

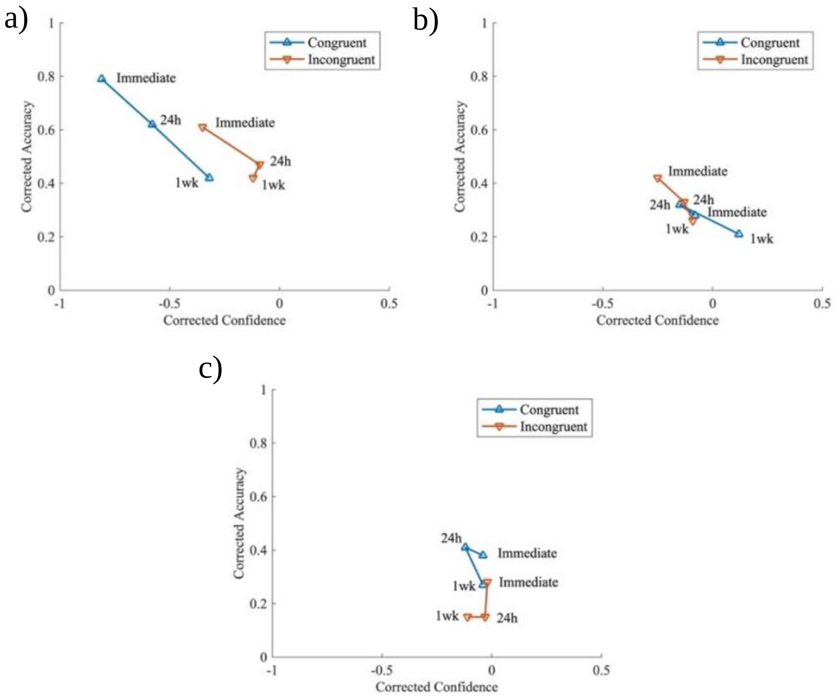

Figure 2.14 Average corrected accuracy and corrected confidence for Congruent and Incongruent targets at each of the three test sessions for the No Change (panel a), Contextualised (panel b) and Decontextualised (panel c) conditions.

2.3.4 Sleep Questionnaires

Descriptive statistics from the sleep questionnaires can be found in Table 2.3. Participant responses on the PSQI, the Epworth Sleepiness Scale and the Stanford Sleepiness Scale demonstrate that participants were not excessively tired at any time point during the study.

Table 2.3 Average responses on the Pittsburgh Sleep Quality Index (PSQI, ranging from 1-21), Epworth Sleepiness Scale (ESS; ranging from 0-24) and the Stanford Sleepiness Scale (SSS; ranging from 1-7). Higher scores on the PSQI indicate poorer sleep quality. Higher scores on the ESS indicate excess sleepiness whilst lower scores on the SSS indicate less feelings of tiredness at that moment.

	N	Score (average)	SD
PSQI	40	5.18	± 1.83
Epworth Sleepiness Scale	40	1.95	± 3.86
Stanford Sleepiness Scale	40	2.14	± .67
SSS Immediate	40	1.95	± .71
SSS 24-hr	40	2.25	± .93
SSS 1-wk	40	2.23	± .97

2.4 Discussion

The current study investigated the effects of context, semantic congruency, and time on visual recognition memory. Participants were shown target objects that were either semantically congruent or semantically incongruent with their background contexts. Targets appeared in one of three context-change conditions, and memory was tested at three time points across one week. The results from the d’ analyses showed the effect of congruency only emerged in two of the three context conditions, supporting the idea that the relationship with the scene affects visual recognition memory. Better performance was observed for targets that were semantically congruent with their background in both the No Change condition and in the Decontextualised condition than when the targets were incongruent with their backgrounds. However, when there was a change in scene-context between study and test, there was no difference in performance between the congruent and incongruent targets. In all cases participants adopted a relatively conservative criterion for congruent targets and a more liberal criterion for incongruent targets. This seems a relatively sensible approach to take in so far as participants may have realised that they could be easily misled in responding old whenever they were presented with a congruent scene. To avoid such mistakes, they adopted a more cautious criterion on congruent trials.

Regarding the confidence ratings the results are not particularly revealing in that in two conditions (i.e., Contextualised and Decontextualised conditions) nothing of any note was obtained. In contrast, in the No Change case the confidence ratings dissociated from the memory data in a rather uninteresting way – whereas confidence grew as time to test increased memory accuracy declined, most likely reflecting a shift in criterion towards a more conservative response bias after a 1-wk delay and for targets that were strongly bound to their scene-context (i.e., congruent targets). The more conservative response bias meant that participants were less likely to rate an item as old, decreasing accuracy and increasing confidence when they did report an item as old. More generally, participants were more confident for FAs than they were for Hits. On these grounds although the confidence ratings in the following experiments (i.e., Chapters 3 and 4) are analysed, reporting of these is included in the Appendix.

2.4.1 Context Reinstatement Aids Object-scene Binding

When the scene context is repeatedly reinstated (i.e., in the No Change condition) and congruent object-scenes are used, tight object-to-scene binding occurs, and the scene facilitates retrieval relative to the incongruent case. For congruent cases, object-to-scene binding occurs automatically (Hayes et al., 2007), forming strong associative links between an item and its contextual details. This is further evidenced by the increased d’ scores in the current study compared to Evans and Wolfe (2022). Keeping the context consistent and presenting the congruent target twice during encoding produces numerically higher d’ scores (Current study: 2.09 vs Evans & Wolfe: 1.62), consistent with the idea that repeated encounters reactivate the original memory trace, leading to reconsolidation of the episode (Zöllner et al., 2023). Memory for the incongruent targets in the No Change condition was statistically worse than for the congruent targets suggesting that tight object-scene binding was most evident in the congruent cases. Furthermore, the repeated presentations at encoding and reinstated context at test also facilitates better memory for semantically congruent targets, consistent with a growing body of literature showing a memory benefit in both verbal and visual memory for congruent targets compared to incongruent targets (e.g., Alejandro et al., 2021; Atienza et al., 2011; Höltje et al., 2019; Packard et al., 2017; Zöllner et al., 2023).

2.4.2 Tight Object-scene Binding Can Hinder Memory

In the Contextualised condition it appears that such tight object-scene binding that occurred from presenting a target in a repeated scene during encoding acted as a hindrance at test. The data in the No change condition suggest that the repetition of the scene-context during encoding facilitated strong object-scene binding. However, at retrieval the scene-context was changed, and memory was severely impacted – the congruence advantage was abolished in this condition. The context at retrieval no longer served as a useful memory cue for the target object. This supports previous work demonstrating a context shift decrement whereby memory decreases when a target is presented in different contexts at study and test (Hayes et al., 2007). Participants would have to rely on memory for the specific item in this condition due to the change in context making the memory trace less accessible, due to a strong integration of the item with the context during encoding (Hanczakowski et al., 2015). It seems therefore that reinstating a context during encoding can strengthen object-scene bindings particularly for congruent cases. Such bindings prove useful when the object-scene context is re-presented at test and therefore aids memory, but such bindings can also prove to be a hindrance when the context is changed at test.

2.4.3 Decontextualisation Leads to a More Abstract Representation

The data in the Decontextualised condition suggest different factors are at play when the target object is never presented in the same scene more than once. In this condition, the target (e.g., lamp 1) was first presented in bedroom 1, then in bedroom 2 and finally at test in bedroom 3. Here it seems that during encoding a more abstract description of the target (a schema) was being stored in memory as opposed to the specific item (e.g., a lamp in a bedroom) and it was this that was operating as a cue for recognition at test rather than specific object-scene bindings. Given the incongruent nature of the object-scene images used for incongruent cases then no such schemas could have been operating here. Since the incongruent targets are not associated with a relevant schema in the scenes used in this study, the new context provides no helpful retrieval cue, allowing the congruency effect to emerge in this condition. Overall, breaking the object-to-scene binding during encoding by presenting the target objects in different background contexts might lead to a more abstract representation of the target, allowing a semantically related new scene context to activate the relevant schemas that would facilitate retrieval. More generally, the results further demonstrate that the relationship of the target object to the scene impacts memory for the targets.

2.4.4 Congruency Effects in Visual Recognition Memory

One of the key remaining questions is whether the congruency effect is carried by processes at encoding or retrieval or at both. A reasonable assumption that arises from the present results is that, in the Decontextualised condition, the congruency effects are driven by retrieval at test. This is evidenced by the fact that the targets were never shown in the same scene, yet congruent targets were better remembered than incongruent targets. Since the scene context was never reinstated, and the object had only ever been presented in two different scenes, the object was not tightly bound to either scene. On the similar assumption that the incongruent objects were also not tightly bound to any particular scene the effect must be driven by processes operating at retrieval. Congruent targets in this condition were better recognised most likely due to the scene providing enough of a semantic context to facilitate retrieval of the gist of the memory trace (e.g., a sofa is highly likely to appear in a living room scene), suggesting that the congruent targets are bound with an abstract scene description which acts as a cue at retrieval. The present results suggest that a more semantic representation of the target object is being stored as opposed to the memory trace being updated via encoding variability and reconsolidation. If encoding variability was leading to a more distributed representation of the target, then recognition should be similar between the congruent and incongruent targets. This is not the case as congruent targets were significantly better remembered in the Decontextualised condition. The inconsistent scene-context present in the incongruent condition was unable to activate related networks at retrieval, leading to a detriment in performance and the emergence of the congruency effect in this condition.

In fact, it is quite possible that different effects of congruence are independently associated with encoding and retrieval, respectively. In the No Change and Contextualised conditions, one could argue that there is tight object-to-scene binding taking place at encoding due to the target being presented in the same background twice. Retrieval in these two conditions is either facilitated by the reinstated context (i.e., No Change condition) or hindered by the changed context (i.e., Contextualised condition). Regarding encoding: it can be argued that incongruent targets challenged schema expectancy upon presentation, weakening the object-scene binding and decreasing recognition memory for those targets. If the semantic context activates the relevant schema shortly after a stimulus is presented (Davenport & Potter, 2004) and works in a way that predicts what should be present in the scene (Oliva & Torralba, 2007), then it is possible that no object scene bindings were established. In other words, because the object was so discordant with the background scene no object-scene bindings were established.

One of the key findings of Evans and Wolfe (2022) was that memory for incongruent targets was better than memory for congruent targets. This was based on the observation of a memory benefit when the target objects were semantically incongruent with the scene and the context changed between study and test (Experiments 7 vs 8). In addition to there being no significant difference between the congruent and incongruent targets in the Contextualised condition, the congruency effect favoured congruent targets in both the No Change and Decontextualised conditions. It is important to note that all the targets in the present study were shown twice during encoding unlike the targets in Evans and Wolfe (2022) which were only shown twice in Experiment 10. Their results showing a benefit of incongruent targets arose from comparing results from two separate experiments, using separate groups of participants, in which targets were presented once during study and were then tested in a second, new scene at test which occurred immediately after encoding. They argued that the incongruent targets provided a release from visual crowding that allowed those targets to be better remembered. Recent research has shown that an increase in the number of presentations during encoding leads to a stronger, more durable memory trace, increasing memory performance and reducing forgetting (Denis et al., 2020, 2021). One could argue that the immediate test procedure used in Evans and Wolfe (2022) did not allow time for consolidation of the targets to take place given that consolidation occurs over time (Dudai, 2004; Roesler & McGaugh, 2019) often after a period of sleep (e.g., Fogel et al., 2009; Nadel et al., 2012; Yonelinas et al., 2019; Zhang, 2022). It is therefore plausible that the repeated presentations in the present study facilitated the formation of associations between the target objects and the scenes, favouring the integration of congruent targets into a cohesive memory trace. Since there was no time for consolidation to occur in the studies conducted by Evans and Wolfe (2022), the incongruent targets potentially remained novel and anomalous with the scene context, leading to an increase in short-term recognition of those targets.

It appears therefore that the critical difference between the experiments reported here and those reported by Evans and Wolfe (2022) is whether or not the same target appeared twice in the same scene during encoding. Clearly a within-participants design could be used to examine this issue, but this is something that remains open for future research.

2.4.5 Recognition Memory Decreases Across the Week

Unsurprisingly, recognition performance was highest when testing occurred immediately following encoding and lowest after a 1-wk delay in all three context conditions. As the length of time between study and test increases, opportunity for forgetting increases. It has been postulated that immediately testing memory for some of the items after encoding leads to interference in later consolidation of the memory traces (Yonelinas et al., 2019). This is due to a high contextual overlap between the items being retrieved from the previously encoded memory traces. Further, retrieval induced forgetting occurs when similar but competing memory traces are retrieved, leading to interference of the related memory traces that are not retrieved (Anderson & Hulbert, 2020), particularly when they are semantically related (Maxcey et al., 2018). Newly encoded information must undergo consolidation in order for the fragile memory traces to become more stable in long-term memory (Hupbach et al., 2007; Nadel et al., 2012; Roesler & McGaugh, 2019; Tambini & D’Esposito, 2020; Winocur & Moscovitch, 2011). It is possible that our immediate testing of a subset of the targets interfered with memory consolidation by reactivating some of the previously encoded traces. Our targets were presented in real-world indoor scenes, thus creating contextual overlap (e.g., objects appeared in various kitchens, bedrooms, living rooms, etc). This could have interfered with the memory traces of similar targets, especially when related targets were retrieved at the first test. Further, whilst studies have shown that briefly presented images can be retained for long periods of time, they are likely to be forgotten if the recognition task contains highly similar foils (Potter et al., 2002). The foils used in our experiment belonged to the same categories as the targets (e.g., chairs, tables, kayaks, machinery), and were placed in related scene contexts (e.g., living rooms, bedrooms, kitchens). It is possible that the high similarity between our targets and foils increased false alarms and led to forgetting. Future research should consider using a more varied range of categories to reduce overlap between targets and foils.

It is also unsurprising that recognition performance was lowest after a delay of one week. Detailed representations of visual memory traces persist short-term, especially when a large number of images have been encoded, but the details of the episodes decrease with time (e.g., Andermane & Bowers, 2015; Brady et al., 2013). Forgetting across the week is additionally pronounced when the tested object is in a different position or arrangement (Mercer & Jones, 2019), a crucial aspect not controlled for in the current study. Since decontextualisation occurs gradually over time and works in a manner that removes the contextual details of the episode, it can be assumed that the targets in our study were more generalised at the 1-wk test which led to a decline in performance after the long delay. Additionally, our target objects in the Contextualised and Decontextualised conditions were not only encountered in a different background context, but also from a different perspective. The lack of details retained after a week combined with viewing the target from a different perspective could explain why we observed a linear decrease in recognition memory across the week.

2.5 Conclusion

The present study examined how semantic congruency and context impact recognition for elements of visual episodic events. The strength of binding of elements in a visual episodic event contributes to memory for individual elements. When the same object-scene image is presented twice during encoding and at test then congruent targets were more accurately remembered than incongruent targets. Further evidence, however, suggests that different effects operate at encoding and test, respectfully. It is suggested that tight object-scene binding for congruent targets comes about when the scene context is repeated during encoding (and when that context is reinstated at test). Under these conditions memory for the target objects is facilitated as shown by the congruency memory advantage. However, when for congruent targets, the scene context is repeated during encoding but changes between study and test, then memory suffers: the congruency memory advantage was abolished. Furthermore, when a congruent target appeared in different scene contexts of the same type (e.g., lamp in bedroom) at both study and test, the memory trace becomes more gist-like and schematic in nature. In this condition, the congruency memory advantage re-emerged. Here the scene at test matches with the stored gist representation and this benefits memory. Such a process does not occur for incongruent targets because no gist “cues” are generated during encoding and therefore these do not operate at test.

This general framework for thinking is explored further in the ensuing experiments and an attempt is made to garner further support for the idea that different effects of congruency operate, respectively at encoding and test.

3. Verbal Episodic Memory and Context

3.1 Introduction

The results from Chapter 2 demonstrated that visual episodic memory was sensitive to congruency effects in several ways. Firstly, semantically congruent targets were recognised better than incongruent targets when the target object was encountered in a repeated scene-context during encoding and in the same scene at test. An implication being that under these conditions congruent targets are tightly bound to their scene context. However, the tight object-scene binding that resulted from seeing the target object in a repeated scene during encoding hindered memory when the scene-context changed at test, abolishing any benefit of congruency. It remains unclear whether these effects are carried by processes operating during encoding and/or retrieval. Finally, when the target object was encountered in three, unique scenes (i.e., two scenes during encoding and a third scene at test), a congruence advantage again emerged. In this case, a schematic representation of the object in the type of scene was retained in memory (e.g., a toaster in a kitchen rather than the specific toaster in a specific kitchen). Under these conditions the new context served as a useful retrieval cue for the congruent targets, whilst providing no such cue in the incongruent cases. Within this condition, it is assumed that processes at retrieval drive the congruency effects, since no object-scene binding occurs during encoding.

Context effects have been extensively observed within the memory literature and there is a strong body of literature that supports sensitivity to congruency in the verbal domain. It remains unclear whether context and congruency affects verbal episodic memory to the same extent that it influences memory for visual events. The two domains have been speculated to share processing characteristics (Shepard, 1967) which suggests that they might also share encoding and consolidation processes. The present research aim was to determine whether the congruency and context effects observed during retrieval in the previous chapter map onto those observed in verbal episodic memory. It has been argued that processing of visual stimuli involves verbalising the event (Baddeley, 1999) and that verbal episodic memory follows the same capacity rules that visual episodic memory follows (Standing, 1973). Standing (1973) demonstrated that both visual and verbal episodic memory are subject to a decline in accuracy when the number of items to be retrieved increases. However, it has been argued that memory in the visual domain is quantitatively better than memory in the verbal domain (Konkle et al., 2010; Standing, 1973). This is thought to be due to pictures providing a more distinct conceptual trace (Konkle et al., 2010). For example, if one is shown a picture of a dog then information pertaining to colour, size, breed, etc is present in the image. Conversely, if one is presented with the word ‘dog’, there is a variety of mental concepts that could come to mind which are unique to the individual and their experiences. Despite there being quantitative differences between the two domains, there is extensive evidence that demonstrates a robust capacity of verbal episodic memory (DeLong et al., 2005; DeLong & Kutas, 2020; Hovhannisyan et al., 2021; Mak et al., 2021; Mak & Twitchell, 2020).

One widely studied aspect of verbal episodic memory is sensitivity to semantic context. The semantic distinctiveness model posits that when a word is encountered in a variety of contexts, those encounters are stored as part of the word’s ‘contextual history’ (Jones et al., 2012; Mak et al., 2021). Each time the word is encountered in a new context, updating of the memory vector occurs and makes the word less context-dependent (Mak et al., 2021). Jones et al. (2012) suggest that encountering a familiar word in a novel context leads to a stronger encoding of that word. Further, they suggest that encoding strength is determined by the amount of overlap between the present context and the contexts stored in one’s lexicon. Therefore, if the context at encoding differs from the contexts that are already present in the current contextual history, stronger encoding is necessary to ensure the word is remembered in association with the new context (Mak et al., 2021). However, it is important to note that each participant will bring a unique contextual history to the experiment which can be due to a variety of factors (e.g., education level; Mak et al., 2021). This could potentially modulate the effects of context on verbal episodic memory.

However, early studies suggest that context reinstatement is crucial for accurate recognition in the verbal domain, in line with patterns found in the visual domain. For example, Light and Carter-Sobell (1970) created a semantic context by pairing nouns with adjectives. They found that recognition memory was highest when the adjective paired with the noun at test was the same as the one paired with it at encoding. They argued that changing the adjective changes the semantic interpretation of the target, disrupting access to the encoding event. More recently, studies have examined whether words that are typically encountered in a variety of contexts are less context-dependent (Mak & Twitchell, 2020). For example, Mak and Twitchell (2020) found across three experiments that words that already have a rich contextual history (i.e., previously encountered in a variety of contexts) are easier to associate with a new context. It can be argued that these words are less context-dependent, making them easier to associate with new contexts.

Another potential reason context places such a strong influence on verbal episodic memory could be due to words having differing meanings based on the surrounding context (Thornhill & Van Petten, 2012). It has been well established that sentence processing is an incremental process whereby each word sequentially builds the context of the episode. This helps to resolve ambiguity of meaning so that memory traces can be rapidly formed (DeLong et al., 2005). Therefore, the entirety of the sentence context is necessary for comprehension, particularly when words have a variety of meanings. Furthermore, it has been proposed that verbal targets, in combination with their contextual surroundings in a sentence, not only activates semantic memory but also an individual's previous experiences (DeLong et al., 2005) in order to resolve the ambiguity that might be present.

However, context is not the only factor impacting verbal episodic memory. The influence of semantic congruency on verbal episodic memory has been widely studied, with behavioural performance demonstrating that accuracy and recognition scores are higher for semantically congruent targets (Bein et al., 2015; Höltje et al., 2019; Packard et al., 2017; Tibon et al., 2017; van Kesteren et al., 2020) than incongruent targets but neural responses are stronger for semantically incongruent targets (Alejandro et al., 2021; Hamm et al., 2002; Höltje et al., 2019; Kutas & Hillyard, 1980; Lau et al., 2011; Schneider & Maguire, 2018) than congruent targets. On the one hand the behavioural evidence appears to suggest the ease of memorising congruent targets whereas, on the other, the neural data suggest difficulties in memorising incongruent targets. For instance, it is possible that semantically congruent words are easier to integrate into coherent memory traces (Lau et al., 2011) than are incongruent words but that additional processing is needed in order to overcome the violations of expectancy (Hamm et al., 2002) when dealing with incongruent targets. Alternatively, it could be that incongruent targets place a more demanding cognitive load on the individual as they attempt to form associative links between a target and its context, evidenced by longer reaction times and increased neural responses in frontoparietal regions when trying to retrieve incongruent, compared to congruent, associative pairs (Gerver et al., 2020).

Researchers suggest that visual and verbal stimuli may activate common networks during processing and retrieval (Hamm et al., 2002). If the two domains activate common networks during retrieval, then one could assume that performance on a verbal episodic memory task will be similar to performance on a similar visual episodic memory task. Therefore, the current study sought to replicate and extend the findings of Chapter 2 within the verbal domain. The role of context and congruency were examined using the same design as previous chapter in order to understand whether the congruency effects observed previously within the three context conditions within the visual domain extend to verbal episodic memory. In the present study, instead of using target objects within scene contexts, semantically congruent and semantically incongruent target words within sentence contexts were used. All target words appeared in one of three context conditions during study and test phases (i.e., No Change, Contextualised, Decontextualised), in order to ensure consistency between the previous visual experiment and the current verbal experiment.

3.2 Methods

3.2.1 Participants

We collected data from 26 healthy adults aged 18-30 years (23 Female). One participant was excluded due to not returning for the 1wk follow-up session. This left us with a final sample of 25 participants. All participants were native English speakers, had normal or corrected-to-normal vision, had no special visual characteristics (e.g., colour blindness or astigmatism), and had no history of drug or alcohol abuse. They had no history of hearing problems, dyslexia, or neurological, psychiatric or sleep disorders, and they did not have a job that involved alternating shift patterns (i.e., day and night shifts). All participants provided informed consent and filled in demographic questionnaires before the study. The current study was approved by the Research Ethics Committees of the Department of Psychology, University of York.

3.2.2 Materials

3.2.2.1 Stimuli

Stimuli consisted of 648 sentences (targets and foils) with paired audio recordings of each sentence. Target words can be found in Appendix 3.A. Sentences were created such that each word appeared either in a semantically congruent (50% of targets) or semantically incongruent sentence context (50% of targets). Semantically incongruent sentences were normed prior to the study to ensure they were incongruent and not bizarre. To do this, sets of incongruent sentences were divided by category (64 sentences per category) and presented to three independent raters. Participants in the pilot task were asked to rate each sentence on a scale from 1 – Not at all bizarre to 10 – Very bizarre. Sentences that were given a score of 6 or less were used in the study. Incongruent foils were obtained from a previous study (Bloom & Fischler, 1980). Target words and foils were underlined to ensure participants knew which words to remember and judge as old/new.

3.2.2.2 Experimental conditions in the experiment

Consistent with the previous chapters, the experiment was designed to include an initial study phase and three testing phases. Each of the target words, half of which were semantically congruent and half semantically incongruent (see Figure 3.1a), appeared twice in the study phase. The order of the appearance was randomised across the phase and was different for each participant. The targets were randomly assigned to one of the three different testing phases intermixed with foil words embedded as targets in unique sentences. An equal number of target words were assigned to three conditions that were intermixed and their order of appearance was randomized; No Change, Contextualised and Decontextualised conditions (see Figure 3.1b). In the No Change condition, the target appeared in the same sentence both times it was presented during the study phase, as well as in the same sentence during the test phase. Similarly, for the Contextualised condition, the target appeared in the same sentence both times it was presented during the study phase but was placed in a new sentence for the testing phase. Finally, the Decontextualised targets appeared in a new sentence each time it was presented during the study phase and in a third, new sentence during the test phase. Each of the three conditions contain the same number of congruent and incongruent target words, with 36 exemplars in each category for a total of 216 target words to be remembered.

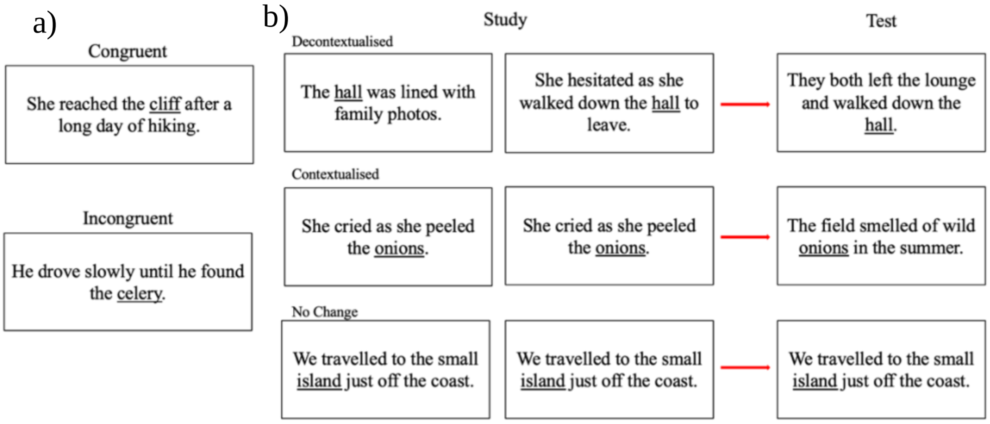

Figure 3.1 a) Example of semantically congruent and semantically incongruent stimuli. Target words were underlined. b) Example of the three context conditions. In the Decontextualised condition, the target word appeared in a new sentence context each time it was encountered. The Contextualised targets appeared in one context at study and in a new context at test. The target words in the No Change condition always appeared in the same sentence context.

3.2.2.3 Experimental Design

Consistent with the study presented in Chapter 2, the experiment was conducted using a within-subjects design. The three main factors examined were Context (No Change, Contextualised, Decontextualised), Congruency (Congruent, Incongruent) and Time (Immediate, 24-h Delayed, 1-wk Follow Up). Within each of the three context conditions, there were an equal number of congruent and incongruent target words. Each unique target was tested once during the experiment as was randomly assigned to one of the three testing sessions.

3.2.3 Procedure

The experimental timeline was identical to the timeline in Chapter 2 with the exception being that the study was conducted in-person rather than online. The experiment was coded in Visual Studio Code using JavaScript, HTML, C# and CSS. It was hosted on Microsoft Azure and run in the lab. The experimental code is available on GitHub (https://github.com/evmadden/SDCexp1). The experiment was conducted on a Windows PC via an experimental website in Google Chrome. Audio recordings of the sentences were presented to the participants via headphones and the volume was set to 25 for all participants.

During the study phase, sentences were presented in the centre of the screen for 4 seconds each accompanied by an audio recording of the sentences. Audio recordings accompanied the visual presentation of the sentences to ensure that participants were exposed to the entire sentence and not focused solely on the underlined word, thus establishing a context for the target. Additionally, due to the length of the sentences, presentation time for each target was 1 second longer than in the previous experiment (i.e., stimuli were presented for 4s instead of 3s). Each sentence was followed by a fixation cross for 1.5 seconds. Like before, participants were presented with 432 sentences (216 unique target words, each presented twice either in the same sentence context or another sentence context) identified by the word being underlined. Participants were instructed to only pay attention to the underlined word and memorise it, ignoring the rest of the sentence. Each target word was presented twice during the study phase such that a total of 432 sentences were presented to the participant. The participants were only asked to respond when a number appeared on the screen instead of a sentence to control for sustained attention of the participants. Throughout the study phase, there was a 10% chance that the participant was shown a number instead of a sentence (see Figure 3.2). Any time a number appeared on the screen the participant was asked to press the ‘spacebar’. This was used to check that the participant was paying attention. If the participant did not respond within 5 seconds, the next sentence was presented. At the end of the study phase, they were then prompted to give a reason as to why the attention check(s) failed.

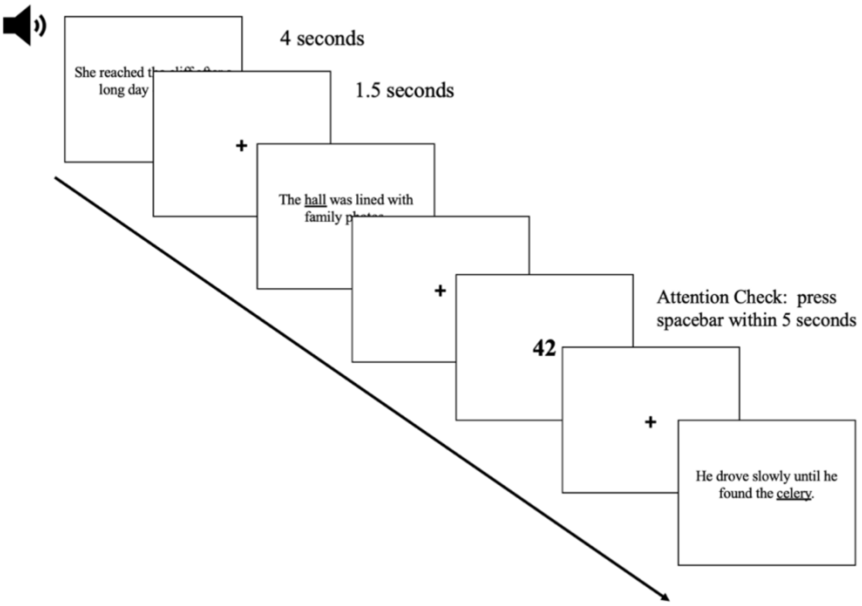

Figure 3.2 Study phase procedure. Participants are shown a sentence for 4s whilst an audio recording of the sentence is played, followed by a 1.5s fixation cross**.** Numbers appear randomly instead of sentences, at which point the participant needs to press the space bar within 5s.

Targets were randomly assigned to one of the three testing phases: Immediate, Delayed and Follow-up. Each testing phase consisted of 72 targets from the study phase plus 72 foils for a total of 144 trials per testing phase. Hence, all of the target words from the study phase were tested during one of the three testing phases, with memory for each target being tested only once during the study. The choice of which target object was tested at each testing phase was randomly assigned for each participant. However, in each testing phase there was an equal number of congruent and incongruent targets as well as an equal number of targets from the three context conditions.

For each test trial, participants were shown a sentence and asked to judge whether the underlined word was ‘old’ or ‘new’ by pressing the left and right arrow keys respectively. Response time for the judgment was recorded in milliseconds. Once the old/new judgment was made, they were asked to rate how confident they were in their answer on a 4-point scale: 1 - not very confident, 2 - somewhat confident, 3 - confident, 4 - very confident. Finally, participants were given feedback on whether their answer was correct or incorrect (See Figure 3.3), with feedback remaining on the screen for 1.5 seconds before progressing to the next test trial. Each test session was self-paced. After the final testing phase, participants were asked to complete the Pittsburgh Sleep Quality Index (Buysse et al., 1989) and the Epworth Sleepiness Scale (Johns, 1991). Finally, participants were debriefed and instructed to contact the researchers if they had any further questions.

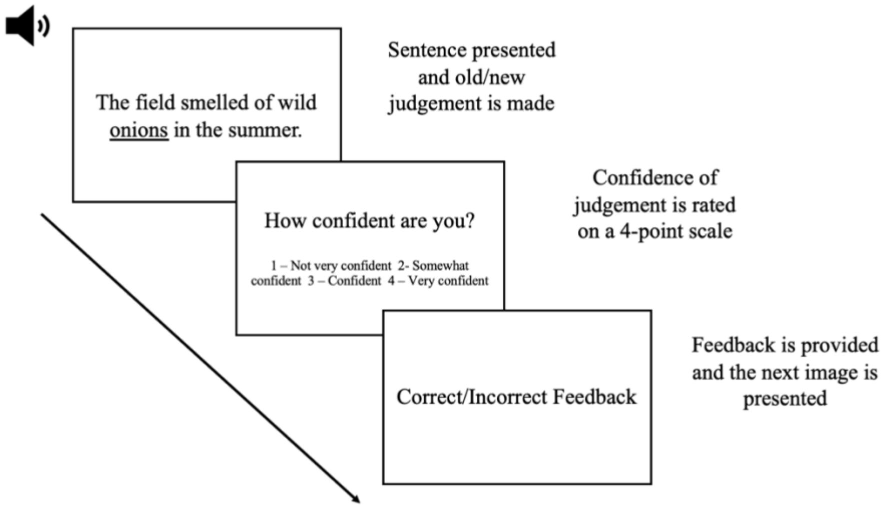

Figure 3.3 Example procedure for each test phase. Participants were shown a sentence whilst a recording of the sentence was played. They were asked to make an old/new judgement for the underlined word as quickly as possible. Next, they rated how confident they were in their answer. Finally, feedback on their answer was given and the next sentence was shown.

3.3 Results

3.3.1 Hits and False Alarms

The first step to understanding the current findings was to check to see whether the results at the Immediate test replicated the pattern of results found in Chapter 2 (see Figure 3.4) by plotting the z-transformed hit rate against the z-transformed false alarm rate (for the comparable figure see Figure 2.7 in Chapter 2). A different pattern of results emerged in the current study (see Table 3.1). In Chapter 2 the pattern of results differed based on whether the targets were congruent or incongruent with the scene context. However, the pattern of results was identical when comparing congruent and incongruent words in sentence contexts. In both cases, there was a decrease in performance when the sentence context changed between study and test, but there was a slight rebound in performance when the word was consistently shown in a new context (i.e., decontextualised). This pattern of results led to the examination whether the differences observed in the data were significant (discussed below).

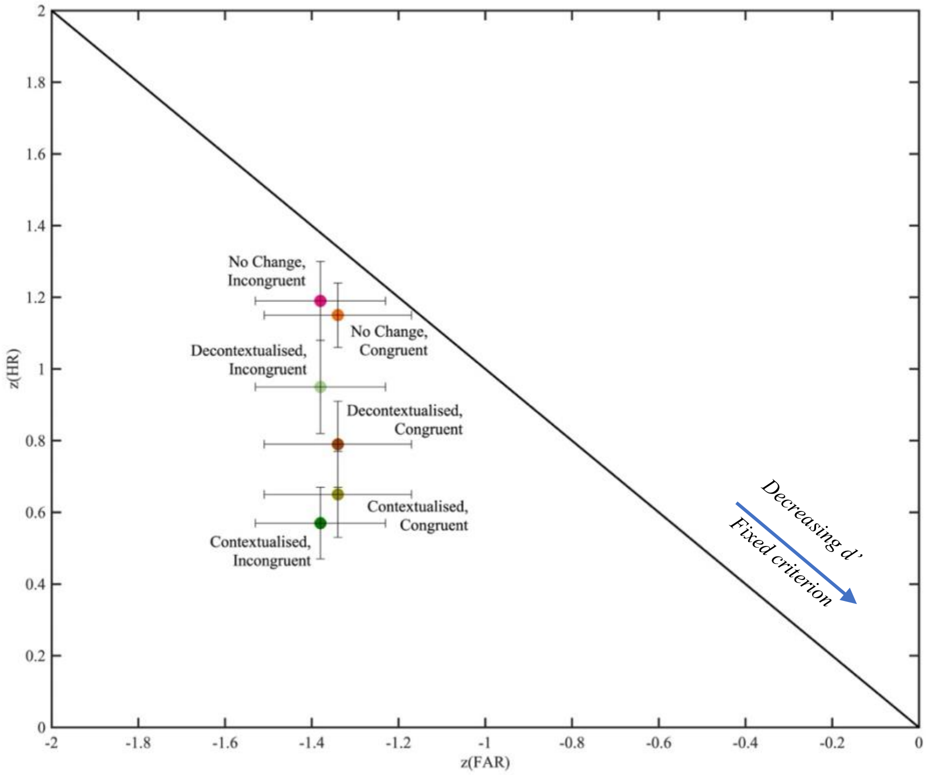

Figure 3.4 Memory performance at the Immediate test represented by z-transformed Hit Rate plotted against z-transformed False Alarm Rate. Better performance is indicated by proximity to the upper left corner and the diagonal black line indicates a neutral criterion. Each data point represents the average for one condition in the experiment. Error bars represent ±1 *SEM*.

Table 3.1 Average d' scores for each condition in the current experiment and from the corresponding conditions in Chapter 2.

Condition	Current Study (d')	Chapter 2 (d')
Congruent		
No Change	2.50	2.09
Contextualised	1.99	.79
Decontextualised	2.14	1.01
Incongruent		
No Change	2.57	1.21
Contextualised	1.95	.82
Decontextualised	2.20	.52

3.3.2 D-Prime and Criterion

In line with the previous study, the foils at test were always embedded in new, unique sentence contexts and different foils were used on the congruent and incongruent trials respectively. As a reminder and consistent with the analyses in Chapter 2, this meant that statistical differences between the three context conditions could not be assessed. The current study followed the same data analysis pipeline as the previous chapter. Three separate 2 x 3 repeated-measures ANOVAs with the factors *Congruency* (Congruent, Incongruent) and *Time* (Immediate, 24-hr, 1-wk) for each of the three context conditions using d' and separately c were conducted. Negative c values indicate a more liberal criterion (i.e., increased likelihood of responding 'old' to a target) whilst positive values indicate a more conservative criterion (i.e., decreased likelihood of responding 'old' to a target).

3.3.2.1 No Change Condition

Consistent with the previous chapter, the effects of congruency and time were first examined on target words that were consistently shown in the same sentence context (i.e., No change condition; context reinstatement) via d'. We observed a main effect of *Time* ($F(2, 48) = 21.41$, $p < .001$, $\eta_p^2 = .47$) but not *Congruency* ($F(1, 24) = .81$, $p = .377$, $\eta_p^2 = .03$). As with the previous study, recognition linearly decreased across the week ($F(1, 24) = 31.54$, $p < .001$, $\eta_p^2 = .56$). However, there was no significant difference between Congruent targets unlike the previous visual stimuli study (See Figure 3.5). Additionally, we failed to observe a significant *Congruency X Time* interaction ($F(2, 48) = .06$, $p = .940$, $\eta_p^2 < .01$).

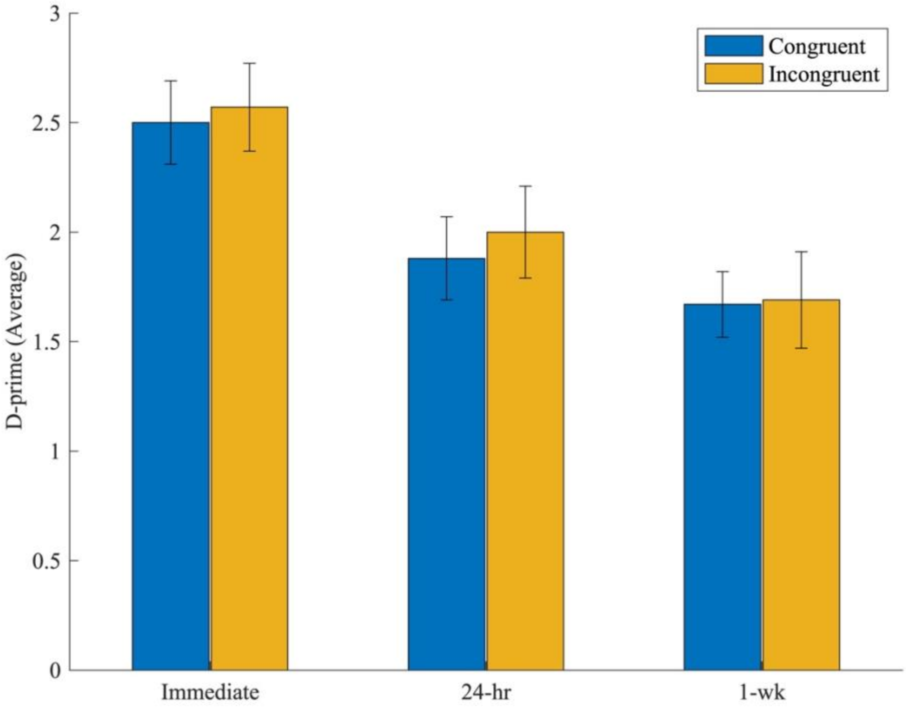

Figure 3.5 Average d’ scores for target memory in the No Change condition at each test session (Immediate, 24-hour delay, 1-week delay) based on congruency (Congruent and Incongruent). Error bars represent ±1 SEM.

Furthermore, when examining the response bias for target words in the No Change condition, there were main effects of *Congruency* ($F(1, 24) = 12.33$, $p = .002$, $\eta_p^2 = .34$) and *Time* ($F(2, 48) = 7.88$, $p = .001$, $\eta_p^2 = .25$). Contrasting with the findings in Chapter 2, participants tended to adopt a more conservative criterion threshold for the Incongruent targets ($M = .02$, $SE = .06$) when compared with the Congruent targets ($M = -.16$, $SE = .06$). Furthermore, participants became more liberal as the length of time between study and test increased. Participants were responding conservatively at the Immediate test ($M = .10$, $SE = .08$), but applied a more liberal criterion at the 24-h ($M = -.10$, $SE = .07$) and 1-wk follow-up tests ($M = -.21$, $SE = .07$). Finally, there was a significant *Congruency X Time* interaction ($F(2, 48) = 3.50$, $p = .038$, $\eta_p^2 = .13$, see Figure 3.6). Simple main effects of interaction showed that the effect of congruency was absent at the Immediate test ($p = .998$), but present at the 24-h ($p = .012$) and 1-wk tests ($p = .002$) and as Figure 3.6 shows this tended to increase as the lag between study and test increased.

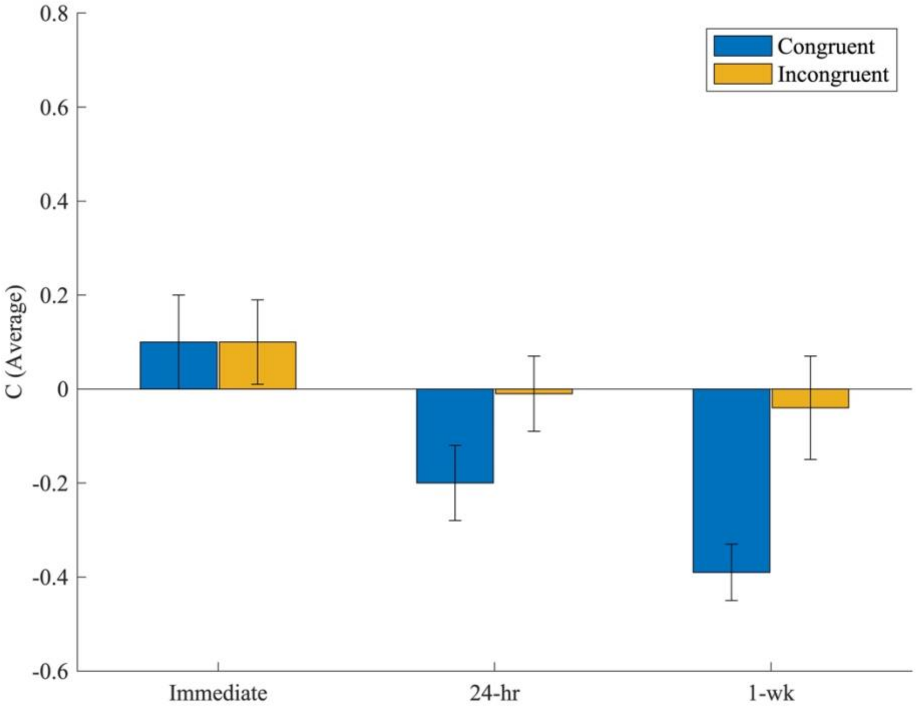

Figure 3.6 Average criterion for old/new decision in the No Change condition at each test session (Immediate, 24-hour delay, 1-week delay) based on congruency (Congruent and Incongruent). Error bars represent ±1 SEM.

3.3.2.2 Contextualised Condition

We next examined the effects of congruency and time on target words that were shown in one context during study and in a new context during test (i.e., Contextualised condition; context shift decrement) based on d'. Again in line with the results from Chapter 2, we observed a main effect of *Time* ($F(2, 48) = 12.26$, $p < .001$, $\eta_p^2 = .34$) but not *Congruency* ($F(1, 24) = 1.52$, $p = .229$, $\eta_p^2 = .06$). Consistent with the results thus far, recognition linearly decreased across the week ($F(1, 24) = 20.72$, $p < .001$, $\eta_p^2 = .46$). There was again no difference between Congruent and Incongruent targets in this condition. For the Contextualised targets, we failed to observe a significant *Congruency X Time* interaction ($F(2, 48) = .63$, $p = .497$, $\eta_p^2 = .03$; see Figure 3.7).

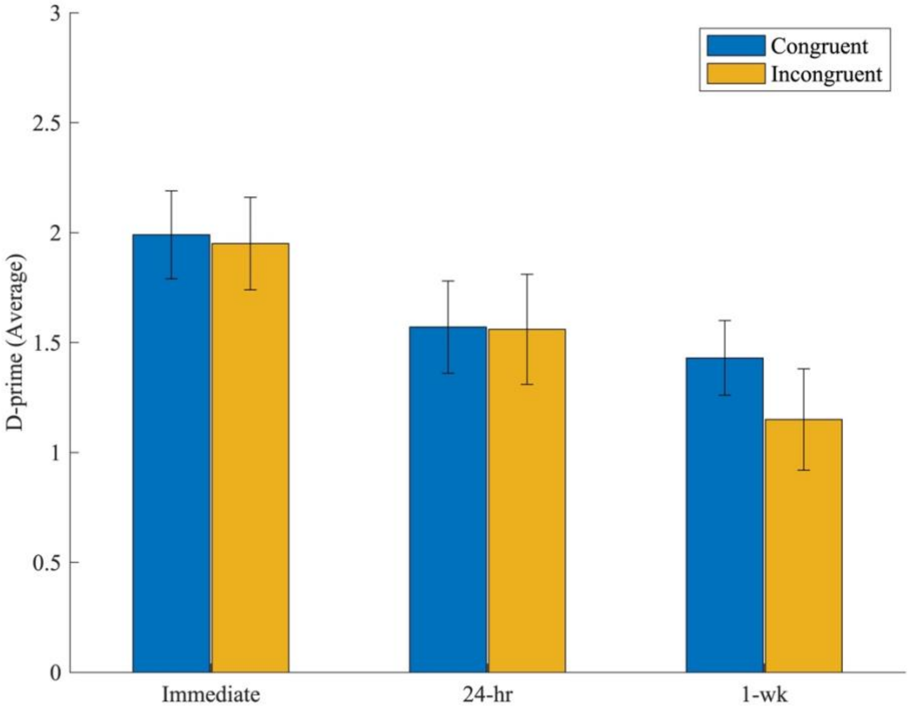

Figure 3.7 Average d' scores for target memory in the Contextualised condition at each test session (Immediate, 24-hour delay, 1-week delay) based on congruency (Congruent and Incongruent). Error bars represent ±1 SEM.

Examination of the criterion threshold for targets whose sentence context changed between study and test revealed there were both main effects of *Congruency* ($F(1, 24) = 21.67$, $p < .001$, $\eta_p^2 = .47$) and *Time* ($F(2, 48) = 16.62$, $p < .001$, $\eta_p^2 = .41$). Participants adopted a conservative decision threshold for both Congruent and Incongruent target words. However, participants were more conservative when making a decision regarding Incongruent targets ($M = .28$, $SE = .05$) compared to Congruent targets ($M = .01$, $SE = .06$). Furthermore, there was a significant *Congruency X Time* interaction ($F(2, 48) = 4.91$, $p = .011$, $\eta_p^2 = .17$; see Figure 3.8). Like with the No Change targets, simple main effects of interaction showed that the effect of congruency was absent at the Immediate test ($p = .596$), but present at the 24-h ($p = .004$) and 1-wk tests ($p < .001$). As Figure 3.8 shows this tended to increase as the lag between study and test increased and was driven by participants switching to a liberal decision criterion for the Congruent targets at the later test sessions.

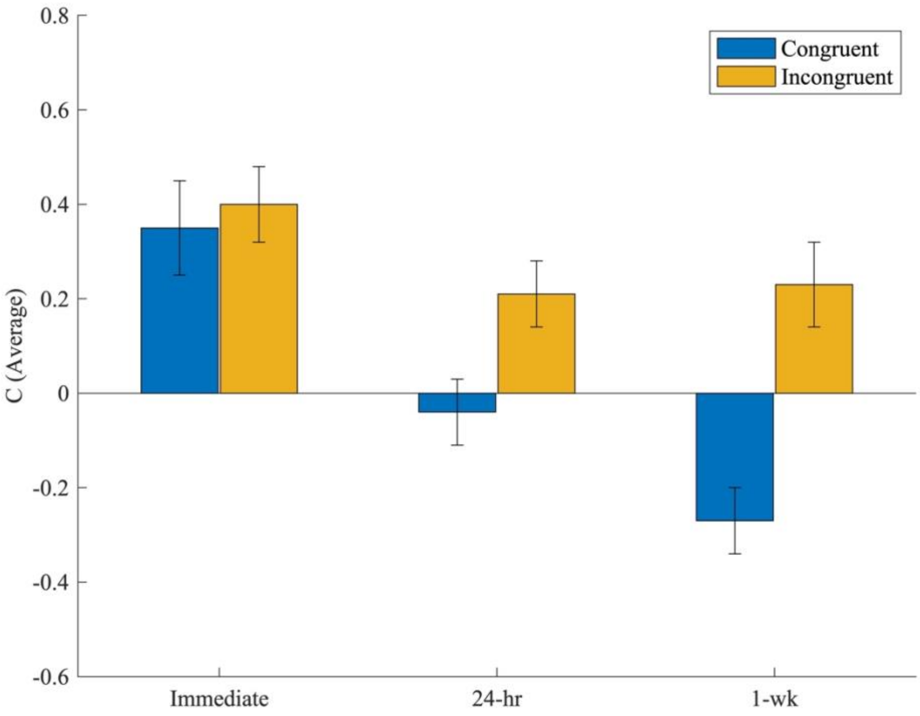

Figure 3.8 Average criterion for old/new decision in the Contextualised condition at each test session (Immediate, 24-hour delay, 1-week delay) based on congruency (Congruent and Incongruent). Error bars represent ±1 SEM.

3.3.2.3 Decontextualised Condition

Finally, we examined the effects of congruency and time on target words that were consistently shown in a new scene context (i.e., Decontextualised condition), based on d’ scores. In this condition, we observed a main effect of *Time* ($F(2, 48) = 8.37$, $p = .001$, $\eta_p^2 = .26$) but again failed to observe a main effect of *Congruency* ($F(1, 24) = .56$, $p = .462$, $\eta_p^2 = .02$). Again, recognition linearly decreased across the week ($F(1, 24) = 20.84$, $p < .001$, $\eta_p^2 = .46$). There was not a significant *Congruency X Time* interaction ($F(2, 48) = .17$, $p = .844$, $\eta_p^2 < .01$; see Figure 3.9), again consistent with the findings from the previous two conditions.

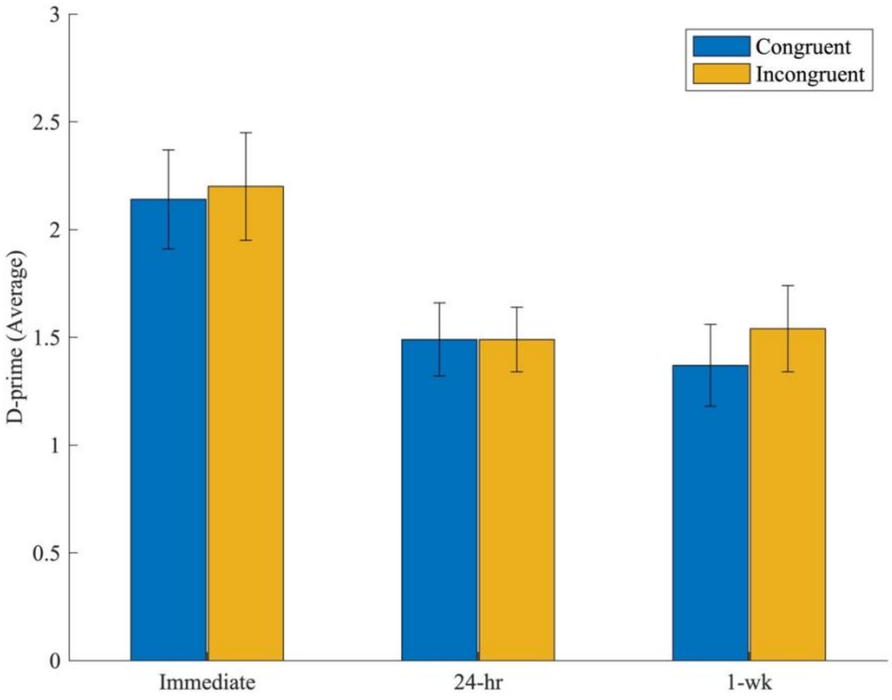

Figure 3.9 Average d' scores for target memory in the Decontextualised condition at each test session (Immediate, 24-hour delay, 1-week delay) based on congruency (Congruent and Incongruent). Error bars represent ±1 SEM.

Furthermore, the results of the criterion analysis for the Decontextualised targets, much like the Contextualised targets, demonstrated there were main effects of *Congruency* ($F(1, 24) = 7.72$, $p = .010$, $\eta_p^2 = .24$) and *Time* ($F(2, 48) = 9.51$, $p < .001$, $\eta_p^2 = .28$). When the targets were presented in continually changing scenes, participants adopted a liberal criterion threshold for the Congruent targets ($M = -.01$, $SE = .05$) but adopted a conservative threshold for the Incongruent targets ($M = .13$, $SE = .06$). Again, the *Congruency X Time* interaction ($F(2, 48) = 3.09$, $p = .054$, $\eta_p^2 = .11$) was found to not be significant (see Figure 3.10) .

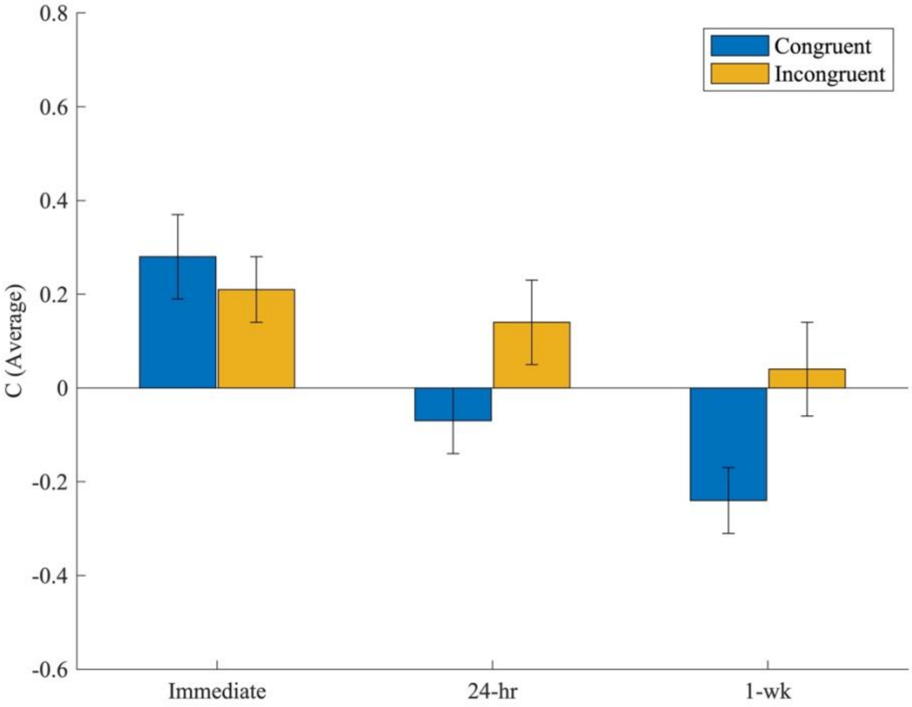

Figure 3.10 Average criterion for old/new decision in the Decontextualised condition at each test session (Immediate, 24-hour delay, 1-week delay) based on congruency (Congruent and Incongruent). Error bars represent ±1 SEM.

3.3.3 Sleep Questionnaires

Descriptive statistics from the sleep questionnaires can be found in Table 3.2. Participant responses on the PSQI, the Epworth Sleepiness Scale and the Stanford Sleepiness Scale, consistent with the previous chapter, demonstrate that participants were not excessively tired at any time point during the study.

Table 3.2 Average responses on the Pittsburgh Sleep Quality Index (PSQI, ranging from 1-21), Epworth Sleepiness Scale (ESS; ranging from 0-24) and the Stanford Sleepiness Scale (SSS; ranging from 1-7). Higher scores on the PSQI indicate poorer sleep quality. Higher scores on the ESS indicate excess sleepiness whilst lower scores on the SSS indicate less feelings of tiredness at that moment.

	N	Score (average)	SD
PSQI	25	5.16	± 1.86
Epworth Sleepiness Scale	25	5.64	± 2.89
Stanford Sleepiness Scale	25	2.23	± .52
SSS Immediate	25	2.24	± .78
SSS 24-hr	25	2.16	± .62
SSS 1-wk	25	2.28	± .84

3.4 Discussion

The present study aimed to replicate the pattern of results from the previous chapter within the verbal domain. The goal of the present experiment was to examine whether the pattern of congruency results observed in different context conditions for visual scenes would be observed for verbal stimuli as well. The same experimental design employed in Chapter 2 was employed in the current study with the exception that target words in sentences were used instead of target objects in real-world indoor scenes. Semantically congruent and incongruent target words were embedded in sentences and presented both visually and auditorily. Target words could appear in one of three context conditions: the same sentence repeated twice during encoding and again at test (i.e., No Change), the same sentence repeated twice during encoding and a new sentence at test (i.e., Contextualised) or two unique sentences during encoding and a third new sentence at test (i.e., Decontextualised). Broadly speaking, the pattern of results from Chapter 2 did not generalise to the verbal domain. In the previous study using target objects in scenes, a congruency advantage was found when the encoding context was reinstated and when the context changed each time the target was encountered. However, in the present study, there were no significant differences in d’ between the congruent and incongruent targets in any of the three context conditions. It is only when comparing the criterion used to make old/new judgements that main effects of congruency are observed. In all three context conditions, the criterion was similar at the immediate test between the two types of targets, but by the 24-h delayed test, participants were adopting a conservative criterion threshold for incongruent targets and a liberal threshold for congruent targets, demonstrating that they were less likely to respond ‘old’ to incongruent targets and were more likely to respond ‘old’ to congruent targets. Such effects are not obviously reflecting aspects of memory – participants appear to be operating with a bias such that they are more accepting of congruent sentences as being more likely to have occurred than incongruent sentences as the delay between presentation and test increased. Nonetheless, there are other effects of memory in the data: overall memory performance decreased as the time between presentation and test increased. It is also notable that memory in the present study was generally good and numerically better than that with visual objects and so the lack of congruency effects here are not a by-product of poor memory for words per se. Taken together with the results from the previous study, the results suggest that congruency operates differently in the verbal and visual episodic memory domains.

There is a large body of literature that has demonstrated a memory advantage for semantically congruent verbal stimuli, with a large number of verbal congruency studies using word pairs or word lists (see Bein et al., 2015; Höltje et al., 2019; Packard et al., 2017; Tibon et al., 2017; van Kesteren et al., 2020): Few studies have used sentences (Besson et al., 1992; Brunellière et al., 2020; Flegal et al., 2014; Haeuser & Kray, 2023; Imai & Richman, 1991; Kutas & Hillyard, 1980; Reggev et al., 2018; Shoben et al., 1978), with most showing a congruency advantage with respect to memory for specific target nouns embedded in the sentences. For example, Besson et al. (1992) presented participants with sentences that ended with either a congruent or incongruent word, with each sentence being presented as one word at a time. Participant memory was tested using a cued-recall paradigm where participants were shown the original sentence and were asked to fill in the final word, with results showing a congruent advantage. For the studies above that employed a recognition task in their paradigms, participants were being tested on aspects such as reading speed and comprehension rather than specific target memory, meaning that the target word memory tests were incidental and unexpected. Additionally, they were only presented with the target word at test rather than the entire sentence. Only two studies have shown an incongruent benefit when the incongruent sentences are intermixed with congruent sentences (Imai & Richman, 1991) or when there are relatively few incongruent trials present during encoding (Reggev et al., 2018). However, none of these studies manipulated the sentence contexts in which a target word was presented (e.g., presenting the target multiple times in different sentence contexts). The current study did not examine memory for the sentences, however by presenting a sentence with the target at test, the paradigm allowed manipulations in context during retrieval. More importantly, participants in the present study were explicitly instructed to only memorise the underlined word. The present results demonstrated no differences between congruent and incongruent targets in any of the three context conditions when targets were presented in sentence contexts but with instructions to only remember the target word. Such testing conditions differ markedly from the other studies on verbal congruency which have shown effects of congruency. It appears that, unless participants are explicitly forced to attend to the entire sentence (e.g., by presenting the sentence word by word), they are able to filter out the unnecessary sentence context, abolishing any effect of congruency. It is due to this lack of difference in congruency that attention is turned towards differences between verbal and visual episodic memory. One possible consideration for future research would be to replicate the current study but present the sentences one word at a time to see whether the classic congruency advantage would emerge.

One possible explanation for the lack of congruency difference is that the sentence contexts provided a means of visualising the event, which could differ from person to person. For example, pictures provide consistent detailed information for each participant (e.g., a specific dog in a specific room) whereas target words in sentences can lead to different mental images for each person (e.g., the word ‘dog’ might be visualised as a Border Collie for one person and as a Great Dane for another). Previous research has demonstrated that pictures are better remembered than verbal stimuli due to the extra details images elicit, narrowing down the genericness of the mental representation that a word elicits (Konkle et al., 2010). The imagery multiprocess model argues that recall of atypical sentences will be enhanced, relative to common sentences, due to participants finding the sentences more interesting and needing increased effort to encode elaboratively (Imai &

Richman, 1991). Participants in the present study were required to form their own mental image of the event due to the verbal nature of the stimuli. Whilst the congruent targets will have been aided by pre-existing schemas, the incongruent sentences may have been more interesting to the participants leading to a more imaginative and elaborate encoding of the event. Thus, both types of target words were deeply encoded, abolishing congruency effects in the present results. This is further supported by the schema-plus-correction model whereby schema-consistent information (i.e., congruent) is not explicitly stored in memory due to the schema facilitating recollection of the event whilst schema-inconsistent information (i.e., incongruent) leads to detailed and vivid representations of the event (Lampinen et al., 2001). Under this model one could assume that the incongruent targets were aided by a detailed, vivid encoding of the sentence whilst the congruent targets were aided by the relevant schemas. This could lead to differences between the visual and verbal episodic memory systems due to the visual pictures providing equally detailed representations for all participants. If this model were assumed, the similar strength of the representations between congruent and incongruent words in sentences abolished any congruency advantage which is in contrast with the increased memory for congruent targets observed in the previous chapter with pictures.

The lack of vivid details in the sentence contexts in the present experiment perhaps leads to the participants having to create their own detailed representations, to which they visualise the congruent and incongruent targets equally. Successful visualisation has been previously shown to significantly improve immediate recognition compared to unsuccessful visualisation (Denis et al., 2020). Recent research has shown that when participants are instructed to create an integrated, vivid mental representation of an associative pair, memory is significantly better than when they are instructed to create separate mental representations for each item in the pair (Denis et al., 2023). Perhaps the sentence contexts, despite being congruent or incongruent, lead to vivid representations of each sentence that aided binding and memory for both types of targets to the same extent. Baddeley (1999) argues that, even when presented with visual stimuli, committing the target to memory involves verbalising the details of the event. However, this would suggest that the congruency advantage present in Chapter 2 should also be present in the current results. It is also highly plausible that the incongruent sentences in the present study created a violation of semantics rather than being truly incongruent like the target object counterparts in Chapter 2. For example, Haeuser and Kray (2023) found that a congruency benefit only emerges when prediction errors (for incongruent contexts) occur early on when reading (e.g., when the incongruent word is early on in the sentence). The incongruent stimuli in the present study were not controlled for where in the sentence the target word appeared, therefore potentially diminishing the congruency differences. Whilst visualisation and semantic violations rather than true incongruence offers an interesting possibility for the lack of congruency differences in the present study, it cannot account for the difference observed between visual and verbal episodic memory.

One final, and most likely explanation, for the lack of congruency differences observed in the present study is due to the way in which participants are able to dissociate the target from the context. Indeed, context facilitates the holistic representation of an event in visual episodic memory (Hollingworth, 2006), and the context serves as one of the critical factors involved in successful retrieval at a later time (Yonelinas et al., 2019). Furthermore, visual episodic memory studies have shown that congruent target objects are automatically bound to their scene-context, regardless of whether the scene is relevant to the task at hand (Hayes et al., 2007). The results from the previous chapter support these findings, suggesting that target objects appearing in scenes provides a holistic representation that is difficult to overcome. However, the present results suggest that participants are able to filter out the unnecessary sentence context and focus solely on the target words for a later memory test. In both studies (i.e., Chapter 2 and Chapter 3), participants were told to focus on the target word/object and to only memorise the target. No congruency advantage emerged in the present study despite using identical contextual manipulations to those employed in Chapter 2 where a congruency advantage emerged. Therefore, one of the key differences between verbal and visual episodic memory is that target objects are difficult to separate from their scene-context whereas target words are more easily dissociable from their sentence contexts.

In an earlier study by Davenport and Potter (2004), evidence was garnered for the idea that objects and scenes are rapidly processed simultaneously after very brief presentations (~80 ms). Arguments about the interactive nature of processing objects and scenes have been made, primarily that, when objects and scenes are presented together, both might be mutually constraining (Davenport & Potter, 2004) in rapid visual tasks. Mutual constraint of having an object in a scene appears to be present in the results of Chapter 2, but not within the confines of the present sentence contexts. It appears to be the case that participants in the present study were able to encode specific target words, successfully ignoring the contextual information as instructed (evidenced by no congruency effects). However, they were unable to filter out the contextual information when presented with objects in scenes (i.e., Chapter 2), suggesting it is not possible to encode an object without also encoding the background details and supporting the argument of a holistic representation of visual events put forth by previous researchers (Davenport & Potter, 2004; Hayes et al., 2007; Hollingworth, 2006; Yonelinas et al., 2019). Detection of scene congruency occurs rapidly, with detection of congruency information occurring after an 80 ms presentation and demonstrating that objects and backgrounds are perceived concurrently when they appear as a complete picture (Davenport & Potter, 2004). The ability to filter out unnecessary and distracting contextual details (as observed in the present experiment), focusing solely on the target word, presents an important distinction between verbal and visual episodic memory. The verbal episodic memory system is able to easily adapt to retrieval needs (i.e., filtering out irrelevant contextual details) whilst the visual episodic memory system appears to be more constrained by the extra details present. It is this distinction that contributes to the rationale for the next thesis experiment.

3.5 Conclusion

The present study examined whether the effects observed in the previous study of visual episodic memory would be observed in verbal episodic memory. A nearly identical experimental paradigm was used for the two studies, with the previous study utilising target objects in pictures (Chapter 2) and the current study using target words in sentences. Previous literature in the verbal episodic memory domain has established a strong memory benefit for targets that are semantically congruent, and the results of Chapter 2 suggested that memory should be better for congruent targets. Numerically, memory was slightly higher for the verbal stimuli than for visual stimuli, and in both studies, memory decreased with time. However, no differences were observed between the Congruent and Incongruent targets in any of the three context conditions in the present study. Several factors may have led to a lack of difference between the two congruency conditions. The pictures presented to participants in the previous chapter provided a more detailed representation of the event, facilitating a holistic representation of the event in which the participants were unable to dissociate the scene from the target. With respect to the verbal targets, participants were able to filter out the unnecessary sentence context, making the context unhelpful during retrieval and abolishing a congruency difference. It is therefore assumed that verbal and visual episodic memory operate differently in response to congruency and contextual manipulations.

Appendix 3.A

Congruent Sentences

No Change

The boy kicked his foot.
She bent and picked a fresh daisy.
The tree was a beautiful shade of green.
He won an award for being the best salesman.
The kids played in the white foam.
The old man was diagnosed with leukaemia.
He wanted a nitrogen powered car.
The young girl was scared of going to the dentist.
The child was vaccinated against measles.
She loved spending the holidays with her cousin.
They drove to their weekly hocky match.
He spent his time researching black holes.
His legs were aching after the long run.
They met every Sunday for a game of football.
She went to get her ear pierced last weekend.
The old man learned to cope with his diabetes.
She thought her cough might be a sign of tuberculosis.
Bananas are rich in potassium.
Her degree was in psychology.
His favourite class was zoology.
The couple repainted the lounge blue.
She fell and hurt her knee.
He was rewarded for being a clever businessman.
The child stayed with her auntie for the weekend.
The computer calculated the time to the nearest microsecond.
The diver feared he would run out of oxygen.
The boy struggled with cell biology.
He loved playing catch with his son in the park.
The mayor threw the first pitch at the baseball game.
The young girl wanted to be a teacher when she grew up.
The vintage car was over a century old.
She was delighted when her pink peony bush bloomed.
He cut his finger making dinner.
Phosphorus is a nutrient used in fertilizer that helps plants grow.
The kids were amazed by the chemistry experiment.
The infant bonded quickly with her mother.

Contextualised

He wore a neck brace after the accident.
She wore a scarf to keep her neck warm.
She always wrote with a purple pen.
The girl preferred the purple towel.
The sick man required a heart transplant.
The old man had to wear a heart monitor.
It was time to add more chlorine to the pool.
Chlorine was added to the water to kill bacteria.
There was a single orchid in the windowsill.
He walked towards the orchid on the table.
The girls won the state volleyball tournament.
The crowd cheered for the volleyball team.
He was vaccinated against malaria before his trip.
Malaria was spreading rapidly through the country.
The kids loved watching The Pink Panther.
The dog carried his pink frisbee around the garden.
He gave his wife a rose on Valentine's Day.
She bought the rose gold watch.

The paper was due in one week.
He was exhausted at the end of the long week.
The young man took his niece to the cinema.
She loved babysitting her niece.
The lily bloomed just in time for Easter.
We bought a lily and put it in a sunny window.
He entered the army as an engineer.
He finally passed his test to become an engineer.
She forgot to pick up aluminium foil at the store.
The brushed aluminium pans were getting old.
Their youngest son was diagnosed with mumps.
He told the doctor that he had mumps as a child.
They couldn't wait to meet their new nephew.
He always went to see his nephew play football.
She was the leading scientist in her field.
She wanted to be a scientist when she grew up.
She topped her salad with sunflower seeds.
A single sunflower grew near the fence.
They went swimming in the nearby lake.
They spent the summer swimming in the nearby lake.
She called a plumber to fix the broken sink.
The plumber came to install our new shower.
The farmer tended the crops over the summer.
The farmer sold his crops at the market every weekend.
He was proud to be her father-in-law.
He quickly bonded with his father-in-law.
The sky was turning grey as the storm drew closer.
She loved how the grey sweater went with any outfit.
The old man was a skilled carpenter.
His grandfather was the carpenter that built their wardrobe.
The sport of the week was badminton.
The girl won the county badminton tournament.
The news station required someone trained in meteorology.
The man accepted a professorship in meteorology.
The boy fell and broke his arm.
The doctor put the cast on the boy's arm.
The blue azalea bush stood out in the garden.
The budding azalea stood out amongst the flowers.
The young girl fought constantly with her brother.
He was always annoyed by his younger brother.
The kids were warned about the dangers of syphilis.
They went to the doctor to get tested for syphilis.
The government was worried about carbon emissions.
He changed the batteries in the carbon dioxide detector.
The boy stood up for his sister.
She always wanted a little sister.
She was accepted into the school of medicine.
His family encouraged him to study medicine.
Their song was the number one hit of the decade.
His car was over a decade old.
A zinc coating was applied to prevent rust.
Zinc is used to protect other metals from corrosion.
They spent the evening reflecting on the past year.
Christmas was their favourite time of year.

Decontextualised

He painted the walls beige.
She replaced the beige carpets.

The old lady bought a beige couch during the sale.
The petunia wilted in the hot sun.
She grew the petunia inside in a pot.
The young girl knelt to smell the petunia.
He hesitated a millisecond too long.
A millisecond before he would have been trapped.
It all takes place within a millisecond.
She pitched the first ball at the softball match.
The girls were the best softball players in their county.
She became the coach of her softball team.
She didn't get along with her mother-in-law.
Her mother-in-law helped her plan the wedding.
His mother-in-law lived far away.
The puppy licked his nose when he smelled the roast.
He pulled out a tissue and blew his nose.
She decided to get her nose pierced.
All the kids got their flu jab this year.
She went to the doctor to get her flu test.
Fewer flu cases were reported due to the use of masks.
They collaborated with a professor from another school.
She was thrilled to be offered the position of professor.
He emailed his professor to ask a question.
The bush bears mauve flowers in the spring.
Her hat was topped with a mauve pom pom.
A mauve dress was delicately hung in the shop window.
The young boy got a new fencing sword for Christmas.
He was excited to meet his new fencing partner.
She drove her son to his fencing lesson.
A yellow daffodil grew in the middle of the field.
The young girl picked a daffodil for her mother.
Daffodil bulbs were planted on either side of the pavement.
They hadn't seen their grandmother in over a fortnight.
The cleaner came to the house every fortnight.
The rubbish was collected every fortnight.
The boy was unable to walk after he had polio.
Most children are vaccinated against polio.
Many people who had polio survived.
She got a carnation boutonniere for her date.
The carnation can be many different colours.
The spy wore a red carnation on his lapel.
The scientist heated magnesium alloys for the experiment.
The farmer tested the soil for magnesium in the spring.
He made a magnesium solution for a later experiment.
He painted a picture of a lone iris in a vase.
The iris thrives in drier climates.
The iris flower gets its name from the Greek word for rainbow.
He was fascinated by the chemistry of fluorine.
Fluorine is a pale greenish-yellow gas with a very sharp smell.
He won a prize for his preparation of fluorine.
The lacrosse club celebrated another successful season.
The lacrosse pitch was in the park behind their house.
He started playing lacrosse from a young age.
The newly qualified nurse was assigned to the cancer ward.
She feared her diagnosis would be cancer.
He agreed to try an experimental cancer drug.

Geranium plants are perfect for those new to plants.
The gardener started his geranium seeds indoors.
The geranium grows in moderate climates.
The couple went bowling on their date.
The kids spent the afternoon at the bowling alley.
The young girl had never been bowling.
She set a reminder to water her pansy once a week.
The pansy flower can be a variety of colours.
A pansy can grow to be 9 inches tall.
He struggled with his physiology module.
Physiology is the study of how living organisms function.
He overslept and was late for his physiology lab.
They started a new project in nuclear physics.
He selected physics as one of his GCSE classes.
They celebrated the acceptance of their new physics publication.
The dandelion was a beautiful shade of yellow.
People mistake the dandelion as a weed.
The young girl blew all the seeds off the dandelion.
They went to watch the tennis match at Wimbledon.
He was looking for a tennis partner ahead of the new season.
The storm destroyed the tennis net.
Sociology explores human behaviour from a global perspective.
Sociology is a broad field that covers a wide range of topics.
Sociology investigates the structure of groups and societies.
Plant science is also known as botany.
The botany lab bought several new microscopes this year.
His role in the botany lab was to keep the plants alive.
They learned about the labourer accident in the mine.
The company was advertising for a seasonal labourer.
His father was a poor farm labourer.
She wanted to talk to a lawyer first.
They agreed to talk to their lawyer in the morning.
He was a skilled lawyer and rarely lost his cases.
In eternity there is indeed something true and sublime.
The new couple vowed to spend eternity together.
She waited for what felt like an eternity.
He died in London of typhoid fever.
The typhoid epidemic ripped through the village.
Typhoid is a bacterial infection caused by Salmonella.
The officers used a violet light to inspect the crime scene.
The priest wore violet vestments during Lent.
The restaurant ordered violet coloured napkins for the tables.
She was the best nurse on the ward.
The nurse prepared the operating room for surgery.
The old man called for the nurse to come help him.
Their lab studied the density of helium.
The balloon was filled with helium.
Helium is present in the atmosphere.
He enjoyed studying astronomy in his free time.
Ancient societies were fascinated by astronomy.
He bought a telescope after becoming interested in astronomy.

Incongruent Sentences

No Change

The house featured a spiral grapefruit in the grand entrance.
She rushed to get the spare lime ready for her guest.
He drove slowly until he found the celery.
The essay was the size of a golf ball.
He published an airplane on the war in an attempt to end the conflict.
The girl took her glacier out of its case.
The window contains high amounts of vitamin K.
It was a surprise that the lemon stack still stood amongst the ruins.
The cellar was planted in nice, neat rows.
She was relieved to finish her first independent scooter.
The pants were damp and dark.
They marched to the beat of the hill.
She enjoyed the bluegrass canyon sound.
They issued a leaflet warning as the storm grew stronger.
There were deep gloves just to the east of their campsite.
The only way to access the beans was from the north-east corner.
She woke early and headed to the cucumber.
She put on her wool typhoon as she walked out the door.
He looked out his window to see smoke coming out of the shirt.
They stored the room so it would last through the winter months.
The woman couldn’t decide which roof to buy.
He loved the rich sound of the lake.
The children splashed and swam in the warm socks all afternoon.
The farmer loaded the dormitory with hay for the animals.
The young boy made the perfect paper farmhouse.
The scene from the skirt was breath-taking.
The buildings were destroyed by a book last year.
She preferred the sound of acoustic stream.
The boy got excited that it was almost piano season.
The peach was dark and damp.
The letter made driving dangerous and difficult.
The brochure was expected to make landfall in a couple of days.
The couple looked for a nice cabbage for their holiday.
The young boy quickly outgrew his cloudy shoes.
The old man paced in the garden behind his kale.
The warm climate made growing cello trees easy.

Contextualised

He was known for selling a basement at the market.
A basement requires deep, well-drained soil and full sun.
The paperback crunched under her feet as she walked to the store.
Cold wind whipped a paperback against his face.
He preferred the stability of a hut.
She heard the squeak of a hut in the back yard.
The young girl was surprised by how the floor tasted.
The boy stood his floor up to look like a small forest.
The young girl spent an hour getting ready in the apple this morning.
The couple wanted the master suite to have its own apple.
She watched for soft areas in the grass where a house might get stuck.
She climbed onto the house and sat on its edge.
The controversial kayak quickly circulated the city.
He quickly skimmed the kayak for important safety information.

All the kids in the class had a Razor shack.
The young child was delighted to get a shack for Christmas.
The fanfare was written for a valley.
He played the valley in the big band.
The man buttoned his new sunshine.
He tossed his dirty sunshine onto the laundry pile.
The old woman knitted her daughter a tornado.
He rose and pulled on a tornado.
She reached the jacket after a long day of hiking.
His gaze lifted to the top of the jacket.
She tugged on the hail of her robe.
The woman unfastened the hail around her waist.
The family pre-booked a hotel to pick them up from the airport.
The hotel driver queued outside the train station.
The young boy got a job delivering the morning car.
The reporter interviewed all the witnesses for the car article.
He was a good biker and just bought his third igloo.
She wanted to fit in with the guys so she bought an igloo.
The photographer’s picture was used on the train cover.
He tossed the train aside and stood.
She stopped at radishes to rest from a full day of driving.
The radishes were several miles up the coast and just outside of the town.
The band was looking to add a cliff player.
The cliff trio wowed the audience with their playing.
The couple loved the vaulted apricot in their new house.
He laid on the carpet and stared at the apricot.
The farmer planted a hall among the other leafy crops.
She was amazed by the variety of the hall at the store.
She descended into the dress to find the ponds and lakes.
They emerged into the basin where the dress floor was a sea of wildflowers.
He started learning the desert from a young age.
The woman was well known for her desert playing.
The kids tried to build a corn in the snow.
He learned to make a corn before his artic adventure.
The couple wanted a cavern at their wedding.
The woman played the cavern to sooth the crying child.
The harp fell swiftly from the tree.
He sat in the shade of a harp tree.
She went to the peas her father took her to as a young girl.
The young man dreamed of owning peas one day.
He watched as the newspaper flickered incessantly.
The newspaper struck a large pitch across the pond.
He found the article soothing as he drifted off to sleep.
The article rumbled in the distance.
It was a small vest where no one lived.
He swam to a vest that was not far away.
The kids sat on the pineapple and opened their presents.
The child left all of the toys on the pineapple.
It was a magazine morning with a high chance of rain.
The magazine sky offered a welcome break from the summer sun.
The young girl wore a colourful fog.
The woman tried on the fog to make sure it fit.
The lightning was too long and needed to be hemmed.
He slid his phone into his lightning pockets.

He left the palace quickly in order to make his connection.
The conductor checked to make sure everyone on the palace had a ticket.
A rope was tied around a sweater nearby to keep the boat close to the shore.
He found a large sweater in the path and moved it out of the way.

Decontextualised

The bulletin on the roof was soothing.
The bulletin kept us from hosting the annual cookout.
She woke to a gentle bulletin on the cold morning.
The bartender made her a violin mojito.
Her favourite was the pink violin.
The violin overpowered all other citrus flavours.
The author's new book quickly climbed to the top of the wagon bestsellers list.
He packed a small wagon into his hand luggage.
She laid in the sun and read a wagon novel.
They knew she could write a motorcycle that would win the contest.
The motorcycle was beautifully written and full of imagery.
She was excited to debut her newest motorcycle at the open mic night.
The villagers tended a small banjo orchard.
The young child picked an banjo from the tree.
The drought decreased the banjo harvest this year.
The woman smiled and pulled off her storm.
He slapped his storm together to remove some of the dirt.
She bought a new storm before the weather turned cold.
The tourists could choose between a Spanish or English version of the ship.
They searched for businesses to advertise in their ship.
The ship contained a comprehensive list of activities in the area.
He couldn't decide between red and green drums.
People came to see the drums growing in the vineyard.
Parents should slice drums before giving them to their young children.
The poem lasted most of the night.
The still air became more charged the closer they got to the center of the poem.
The poem was loud and woke him from his sleep.
She played the ocean solo in the jazz band.
He was a founding member of the ocean quartet.
The ocean is smooth in tone.
The album featured rock solos on every song.
The young child received a rock for their birthday.
He pulled a rock out of his pocket and started playing.
The recipe called for the zest of one trumpet.
She was skilled at making a trumpet drizzle cake.
It was sour like a trumpet.
The girl folded her sleet and put it away.
The man packed sleet for his trip.
The boy got new sleet for summer.
They found an old taxi that belonged to their grandmother in the moving boxes.
The research team submitted their paper to the taxi for publication.
She wrote in her taxi every night before going to bed.
A colouring bus and crayons kept the child busy.
You could always find her reading a bus.
She took a trip to the used bus store to find a new story to read.

Each glass of guitar juice is packed with vitamins.
The guitar fields form a striking feature in the valley.
She didn’t know if she was supposed to eat the guitar seeds or not.
The old woman made a trombone cobbler for her grandkids.
They had trombone sorbet for dessert.
He planted the trombone seed in hopes of growing his own tree.
He thought the viola might give him an energy boost.
Viola and vanilla are the perfect combination of energy and relaxation.
She added a teaspoon of viola juice to the bowl.
The ship carefully navigated the cold water to avoid hitting a belt.
They readied the dog sleds to navigate the slick belt.
The path was dangerous and required them to cross the belt.
Clarinet is cheapest during the summer months.
They had a contest to see who could spit clarinet seeds the farthest.
The clarinet was growing larger every day.
Lots of people auditioned for the cave section.
She transitioned from a closed-hole to an open-hole cave.
The cave section was one of the largest in the band.
The ocean was hidden beneath the novel and the air grew cold.
The novel was thick and made it difficult to see.
A damp novel coated the floor of the forest.
The kids competed to see who could eat their entire cymbals.
The chef squeezed cymbals over the dish to finish it off.
She used fresh cymbals in her margarita.
She sat down and pulled her cyclone off.
He made a ball out of a cyclone.
He wore a warm cyclone while hiking.
The summer sun made her want to take the roof off her mansion.
The mansion was well equipped to handle the muddy terrain.
He rented a mansion for their camping trip.
The sunrise slowly and gently lit up the tie below.
They walked carefully along the edge of the tie.
She stood in awe of the vastness of the tie.
She preferred taking truck exams to multiple choice exams.
The final assessment included a truck on the topic of their choosing.
The exam included a short truck component.
The instructor taught them how to flip a tent in an emergency.
They decided to rent a tent instead of buying one for their trip.
He took his tent out as soon as it was warm enough to be on the water.
She stood next to the river as she sang.
The river played softly in the background.
The man taught river lessons in his spare time.
She pulled on her knit hurricane and prepared to go out in the cold.
The man quietly retrieved his hurricane and left.
He removed his hurricane as he sat down to eat.
She didn’t realise how exhausting it was to shop for a new castle.
The castle rocked slightly with a gust of wind.
Headlights flashed on the curtains briefly and a castle door slammed.
They were told they could use any edition of the required Jeep.
The student was upset when she couldn’t find her Jeep.
She used the mock exam at the end of the Jeep to help her study.
The child gently plucked a wild flute from the bush.

He watered the flute bushes in his garden.
She made a flute shortcake for her birthday.
She always kept a light snow in her bag.
The woman quickly put on her snow and rushed out the door.
The man gave the woman his snow as the sun began to set.
She loved the warm pamphlet and the fragrant flowers.
The bride woke to a pamphlet on her wedding day.
The long days and a bright pamphlet were welcomed in the village.
He loved adding saxophone to his smoothie.
She made a lovely saxophone jam.
My grandmother tended the saxophone bushes.

Appendix 3.B

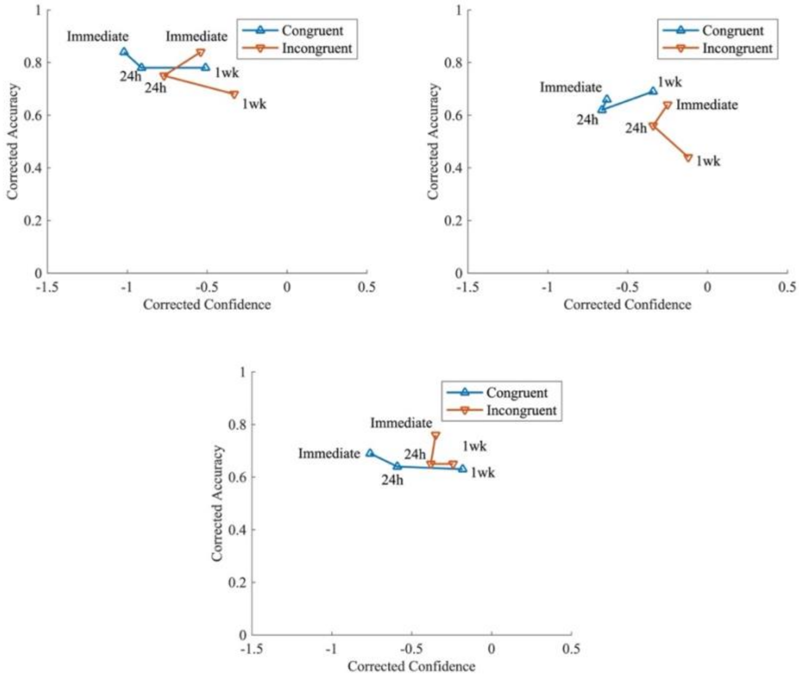

Figure 3.11 Average corrected accuracy and corrected confidence for Congruent and Incongruent target words in sentences at each of the three test sessions for the No Change (panel a), Contextualised (panel b) and Decontextualised condition (panel c).

4. Context, not animacy, determines memory for objects in scenes

4.1 Introduction

Evans and Wolfe (2022) argued that tight object-scene binding occurred in their series of experiments due to the target objects not regularly moving of their own accord, thus reinforcing a “standard representation” of a scene context. They further argued that if an item in the scene did not have a specific place in the scene (i.e., it could appear anywhere), then tight object-scene binding might not take place (Evans & Wolfe, 2022). The results from Chapter 2 demonstrated that, when a target object was semantically congruent with the scene context and the context was repeated, tight object-scene binding facilitated memory for the target object. Additionally, this tight object-scene binding hindered memory when the scene context was changed between study and test. However, presenting the target object in continually changing scenes led to a more gist-like and schematic representation of the target object being stored. This led to superior memory for the congruent cases due to a relevant new scene at retrieval serving as a useful retrieval cue. This raises the question as to what actually happens when targets in the visual recognition paradigm are items that regularly move around (e.g., cats, dogs, etc.) and hence are not normally tightly bound to any particular kind of scene. It is this question that is addressed directly in what follows.

A large portion of visual episodic memory research has used non-living, inanimate targets as stimuli. The functionalist view of memory posits that human memory is fine tuned to solve ‘critical adaptive problems’ pertaining to basic human needs and survival (e.g., securing food/water, recognising and responding to predators (Bonin et al., 2014, 2015; Gelin et al., 2018; Nairne et al., 2013). This has been evidenced by animacy effects in memory whereby animate stimuli (e.g., living things; humans, plants, animals) are remembered better than inanimate stimuli (Bonin et al., 2014, 2015; Gelin et al., 2018; Hovhannisyan et al., 2021; Nairne et al., 2017). It has been argued that due to their link with survival needs, priority is given to processing animate stimuli (Bonin et al., 2014). This is especially evidenced in visual attention and change detection tasks. For instance, New et al. (2007) had participants complete a change detection task using 70 scenes containing animals and non-animals. In their change detection task, participants were asked to detect differences between two rapidly presented, alternating scenes in which one object in the scene was changed. Some trials contained an inanimate object that was changed whilst the other trials contained a change to an animate object. They found that changes to people/animals were not only detected more often, but also more quickly when compared to inanimate objects changing. The animacy advantage observed in both accuracy and reaction time was quite large (21-25% more accurate and 1-2s faster). They argue that the results support the idea that animate targets capture more attention.

Such findings are in line with the associated *animal monitoring hypothesis* (Nairne et al., 2013; New et al., 2007). Increased attention is given to animate stimuli due to ancestorial survival needs, which some researchers have argued has remained a facet of human processing through the generations (New et al., 2007). It has been argued that the human attention system allocates more attention to animals due to animals being able to change behaviour and location rapidly, compared to things like plants which cannot move freely. This makes quick reinspection of the environment crucial for survival and has facilitated the fine-tuning of the attentional system towards goal-directed processes (e.g., survival; New et al., 2007). However, others have argued that animate stimuli are simply more salient and are granted more attention (Nairne et al., 2017; New et al., 2007; VanArsdall et al., 2013). Leading on from the findings that animals attract more attentional resources, research has shown that animacy has been shown to produce robust effects in verbal free recall and recognition tasks, specifically that animate targets (e.g., the names of animals) are remembered better than (names of) inanimate targets (Bonin et al., 2014; Nairne et al., 2013). It has been postulated that correct recognition of animate stimuli is the result of strong contextual binding that takes place during encoding which facilitates better recognition at test (Gelin et al., 2018).

It is important to note, however, that the literature shows a robust animacy effect in verbal long-term memory (LTM) when using words and completing free-recall and recognition tests (Bonin et al., 2014, 2015; Gelin et al., 2018; Nairne et al., 2013; VanArsdall et al., 2013). For example, Bonin et al. (2014) presented participants with two lists of nouns, containing both animate and inanimate words. They were asked to make a categorisation judgement as quickly as possible, labelling each noun as either animate or inanimate. A surprise free-recall task occurred after a five-minute distractor task. Not only were animate words categorised faster, but they were also significantly more accurately recalled than inanimate words. The finding that animate words were better recalled than inanimate words replicated in a subsequent experiment when participants were aware that some form of memory test would occur after encoding (Bonin et al., 2015). Gelin et al. (2018) used a similar experimental design, asking participants to categorise nouns as animate or inanimate followed by a recognition test. They found a strong, significant recognition advantage for animate words compared to inanimate words (*d* = .98), consistent with the findings of Bonin et al. (2014). Furthermore, VanArsdall et al. (2013) paired nonwords with properties that were characteristic of living and non-living things and tested memory for nonwords using both recognition (Experiment 1) and free-recall (Experiment 2) tasks. In both types of tasks, a significant memory benefit for nonwords paired with animate properties was observed. They argue that 'deeper' levels of processing take place for animate stimuli, leading to a memory benefit for animate targets compared to inanimate targets.

Furthermore, Schreiner et al. (2023) recently examined what role animacy plays in binding event elements in verbal episodic memory. During the learning phase, participants were presented with two out of three event elements sequentially (e.g., animal-object, object-location). Participants were asked to imagine the event elements interacting in a meaningful way in order to facilitate vivid imagery of the event. The test phase consisted of presenting a cue word and participants responding whether they remembered the word, knew they had seen it, or did not recognise it. If participants reported having previously encountered the target, they were given a forced-choice recognition task. The authors conclude, across a series of experiments, that animate objects facilitated event binding if the associations during encoding were coherent. They argue that animacy, in verbal studies, facilitates event binding by providing a grammatical subject for representing the event as a sentence. Animate targets provide event structure awareness (Schreiner et al., 2023) which facilitates binding of events into coherent memory traces. Recall back to the findings in Chapter 3 which demonstrated differences between visual and verbal episodic memory. A congruency advantage in memory has been well established in the literature, with the results in Chapter 2 showing a congruency advantage within the visual contextual-binding paradigm. A congruency advantage emerged when the context was consistently reinstated and when the context consistently changed. However, when the same paradigm was used to test memory for target words appearing in sentences, the congruency advantage was abolished in all context conditions, and participants were able to filter out the unnecessary contextual details, focusing solely on the target word. Hence, no congruency effects were observed in the verbal cases. Despite the strong animacy advantage observed in studies using words and its possible faciliatory role in binding verbal events into coherent memory traces, it is still unclear whether this advantage is observed in visual episodic memory as well.

There is a limited number of studies that have looked at animacy effects using images. For example, Sun et al. (2011) asked participants to rapidly categorise 320 objects as animals or non-animals using a go/no-go task. Each object either appeared on a blank background, on a phase-randomised background, or as part of a scene (known as the original scene) and each target was seen three times, once in each background condition. Targets appeared on the screen for 20 ms and participants made a response if an animal was present (no response if it was not). They found that participants were significantly more accurate at responding to animals compared to refraining from making a response when the target was a non-animal in all three context conditions. Accuracy for animate responses ranged from 96.6% to 97.5% (measured by the Hit Rate) whereas accurately refraining from a response when an inanimate was present in the scene ranged from 82.0% to 85.4% (measured by Correct Rejections). In both cases, animate and inanimate items, accuracy was significantly better when the targets appeared as part of a scene context. More interestingly, participants were significantly faster at responding if the animal appeared in its original scene compared to when it appeared in a changed background or in isolation. They argue that the results demonstrate a crucial role of context in in object recognition, however it is also possible that participants were remembering the complete scene rather than the object per se. It is important to note however, their study, whilst using visual stimuli, was not a test of memory.

Limited research has been done to examine whether the animacy effects observed in verbal memory extend to visual memory. In one such memory task utilising images, Bonin et al. (2014) sought to examine whether these effects emerge for pictures. In their second experiment, participants were presented with a series of trials containing a picture of the target paired with the written word for the target (e.g., a picture of a dog appearing on the screen with the word “dog” present on the screen simultaneously). Participants were asked to quickly make an animate-inanimate categorisation for each target object. After a brief distraction tasks, participants received a surprise free-recall task for the names of the pictures they viewed. They found that participants were significantly faster when making categorisation judgements about animate targets and that recall of animate target names was significantly more accurate than for inanimate targets. They compared the findings from their first experiment (using animate and inanimate words) to their experiment using pictures and found that memory performance was higher in the pictures experiment than in the words experiment.

One difficulty here is that it is not clear whether the effects are due to visual memory, verbal memory or both. The stimulus presentation included displaying the word for the target alongside the picture to ensure participants used the correct names for the target pictures. It seems more likely that Bonin et al. (2014) were testing verbal memory via picture presentation rather than visual memory. The presentation of the words presented alongside the pictures combined with a free-recall task for the nouns suggests that the memory measures obtained pertain to verbal memory rather than memory for the target images. Nevertheless, this study provides promising insight into an emergence of animacy effects in visual episodic memory, but limited research has been conducted to date in order to draw firm conclusions. Bonin et al. (2014) argue that animacy effects found in verbal LTM extend to visual LTM. However, to the best of our knowledge, no other research has been conducted examining animacy effects in LTM using pictures, making it unclear what role animacy plays visual LTM.

Evans and Wolfe (2022) argued that objects which can freely move will not be bound to a particular scene context. Hence the corresponding object-scene association will be at best a weak retrieval cue. In contrast targets that are readily associated with particular scenes will generate effective retrieval cues (e.g., a kettle in a kitchen). Given that it is still unclear what animacy effects occur in visual memory, the current study sought to better understand the effect of context and object-scene binding in visual episodic memory using animate and inanimate targets. The results from Chapter 2 demonstrated strong object-scene binding for the semantically congruent targets when the scene-context was consistently reinstated, supporting memory for the target objects. In order to determine if binding with a scene context takes places for targets that can freely move, testing the idea proposed by Evans and Wolfe (2022), the present study only utilised animate and inanimate targets that were semantically congruent with the scene-context. The current study was a partial replication and extension of the study presented in Chapter 2 and utilised the same experimental design. Participants were presented with target objects embedded in real-world indoor scenes during a study phase and recognition memory for those objects was tested three times across one week. Our stimuli consisted of two types of targets: animate objects (e.g., dogs, cats, humans) and inanimate objects (e.g., sofas, tables, beds). The inanimate stimuli in the present experiment are the same stimuli used in Chapter 2 for the congruent targets. This provides an opportunity for replication and to explore whether tight object-scene binding occurs for animate targets.

4.2 Methods

The methods for the current study are identical to the study presented in Chapter 2. The only difference between the two studies were the targets we used.

4.2.1 Participants

We collected data from 48 healthy adults aged 18-30 years (29 Female). Three participants were excluded due to technical difficulties, four were excluded due to missing a testing session and one was excluded for failing the attention checks. This left us with a final sample of 40 participants. All participants had normal or corrected-to-normal vision, no special visual characteristics (e.g., colour blindness or astigmatism), and had no history of drug or alcohol abuse. They had no history of hearing problems, dyslexia, or neurological, psychiatric or sleep disorders, and they did not have a job that involved alternating shift patterns (i.e., day and night shifts). All participants provided informed consent and filled in demographic questionnaires before the study. The current study was approved by the Research Ethics Committees of the Department of Psychology, University of York.

4.2.2 Materials

4.2.2.1 Images

Stimuli consisted of 648 (targets and foils) unique, coloured images gathered via Google Image searches. Target objects were embedded in real-world indoor scene background contexts (e.g., living rooms, kitchens, bedrooms) and identified using a red bounding box. The 216 target objects were either an animate object (e.g., a dog or person; 50% of targets) or an inanimate object (e.g., a couch or bed; 50% of targets). Figure 4.1 provides an example of our animate and inanimate stimuli and Table 4.1 outlines the categories included in our stimuli set for each type of target. Each animate and inanimate category contained 36 target object exemplars and there were 216 object foils (108 animate and 108 inanimate). We are unable to specify the degrees of visual angle or the viewing distance for each participant due to the study being conducted online.

Table 4.1 Exemplar categories for inanimate and animate target objects.

Inanimate Object Categories	Animate Object Categories
Appliances	Bird
Chairs	Dogs
Sofas	Humans
Beds	Insects
Cabinets	Reptiles
Desks	Small Mammals
Lamps	
Chandeliers	
Tables	
Ornaments	

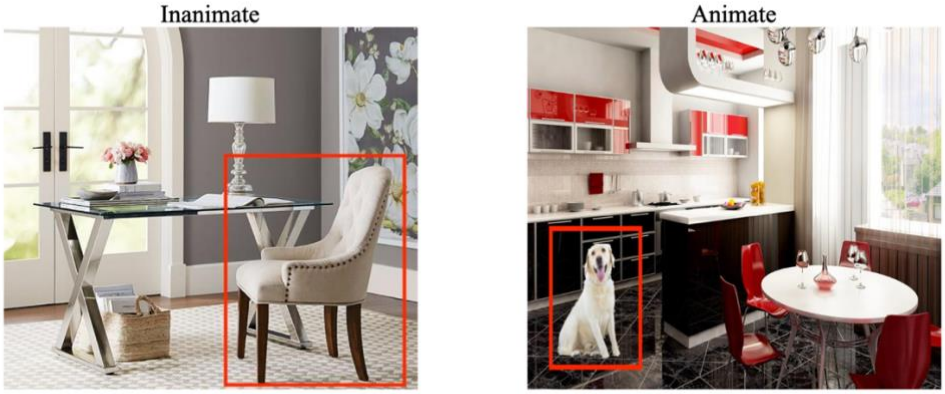

Figure 4.1 Example of inanimate and animate targets within indoor scenes.

4.2.2.2 Apparatus

The experimental apparatus was identical to the one used in Chapter 2. The experimental code for the current study and stimuli are available on GitHub (https://github.com/evmadden/SDCexp1).

4.2.2.3 Experimental Design

Consistent with the previous two studies, the experiment was conducted using a within-subjects design. The three main factors examined were Context (No Change, Contextualised, Decontextualised), Animacy (Inanimate, Animate) and Time (Immediate, 24-h Delayed, 1-wk Follow Up). Within each of the three context conditions, there were an equal number of animate and inanimate target objects. Each unique target was tested once during the experiment as was randomly assigned to one of the three testing sessions.

4.2.3 Procedure

The current study was conducted online and followed the same procedure as the previous experiments (see Chapter 2).

4.3 Results

4.3.1 Hits and False Alarms

The first step to understanding the current findings was to check to see whether the results were similar to those in Chapter 2 (see Figure 4.2 and Table 4.2) by again plotting the z-transformed hit rate (HR) against the z-transformed false alarm rate (FAR; for the comparable figure see Figure 2.7 in Chapter 2). The pattern of results replicates the pattern of results observed in Chapter 2. Like before, performance is best when the target appears in a consistent scene each time it is encountered (regardless of whether the target is an animate or inanimate object) and changing the context leads to a steep decline in performance consistent with the classic context reinstatement and context shift decrement effects. The study reported in Chapter 2 showed a different pattern of results based on congruency when examining the HR and FAR in the Contextualised and Decontextualised conditions, with better performance in the Decontextualised condition compared to the Contextualised condition for congruent targets, and the opposite relationship for the incongruent targets. Unlike in the previous chapter, performance was better, regardless of whether the target was animate or inanimate, when the scene context is continually changing compared to when the target was studied in one context and tested in a new context. This finding is consistent with the idea of decontextualising the target, suggesting that it was less dependent on the context at test due to never appearing in the same scene context. This pattern of results led to the examination of whether the differences observed in the data were significant (discussed below).

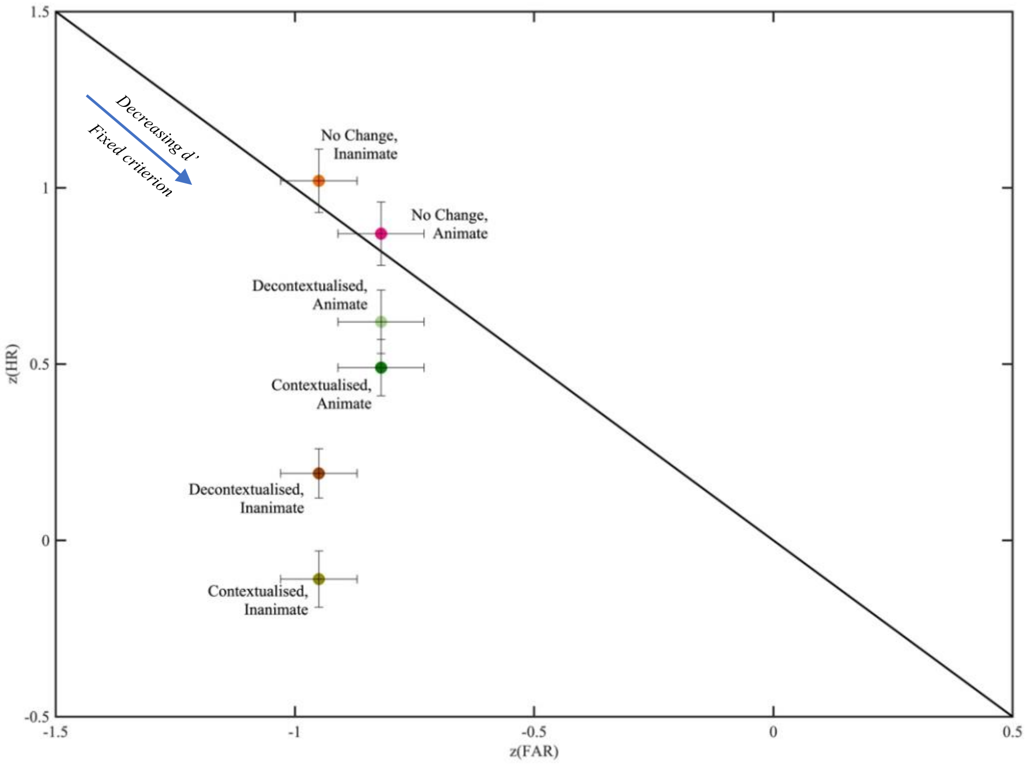

Figure 4.2 Memory performance at the Immediate test represented by z-transformed Hit Rate plotted against z-transformed False Alarm Rate. Better performance is indicated by proximity to the upper left corner and the diagonal black line indicates a neutral criterion. Each data point represents the average for one condition in the experiment. Error bars represent ±1 *SEM*.

Table 4.2 Average d-prime (d') scores at the Immediate test for each condition in the current experiment and from the corresponding condition in Chapter 2.

Condition	Chapter 2 (d')	Condition	Current Study (d')
Congruent		Inanimate	
No Change	2.09	No Change	1.97
Contextualised	.79	Contextualised	.84
Decontextualised	1.01	Decontextualised	1.13
Incongruent		Animate	
No Change	1.21	No Change	1.69
Contextualised	.82	Contextualised	1.31
Decontextualised	.52	Decontextualised	1.43

Note. The stimuli used in the Congruent condition from Chapter 2 and in the Inanimate condition in the present study are the same stimuli.

4.3.2 D-Prime and Criterion

As before, we calculated d' and criterion (c) for animate and inanimate targets in each of the three context conditions and across the three test sessions. Like with the congruency targets, the foils consisted of new animate and inanimate targets embedded in new, unique background scenes. This meant that d' scores were calculated based on animacy and statistical comparisons between the three context conditions could not be assessed. We conducted three separate 2 x 3 repeated-measures ANOVAs with the factors *Animacy* (Inanimate, Animate) and *Time* (Immediate, 24-hr, 1-wk) for each of the three context conditions via d'. Response bias (criterion; c) was calculated and entered into three 2 x 3 repeated measures ANOVAs to determine how participants were adjusting their decision criterion at test using the factors *Animacy* (Inanimate, Animate) and *Time* (Immediate, 24-h, 1-wk). Negative values indicate a more liberal criterion (i.e., increased likelihood of responding 'old' to a target) whilst positive values indicate a more conservative criterion (i.e., decreased likelihood of responding 'old' to a target).

4.3.2.1 No Change Condition

We first examined the effects of animacy and time on targets that were consistently shown in the same scene context (i.e., No Change condition; context reinstatement) based on d'. We observed main effects of *Time* ($F(2, 78) = 56.32$, $p < .001$, $\eta_p^2 = .59$) and *Animacy* ($F(1, 39) = 17.19$, $p < .001$, $\eta_p^2 = .31$). Recognition linearly decreased across the week ($F(1, 39) = 115.27$, $p < .001$, $\eta_p^2 = .75$) and Inanimate targets were significantly better remembered than Animate targets. We failed to observe a significant *Animacy X Time* interaction ($F(2, 78) = 2.35$, $p = .103$, $\eta_p^2 = .06$), demonstrated in Figure 4.3.

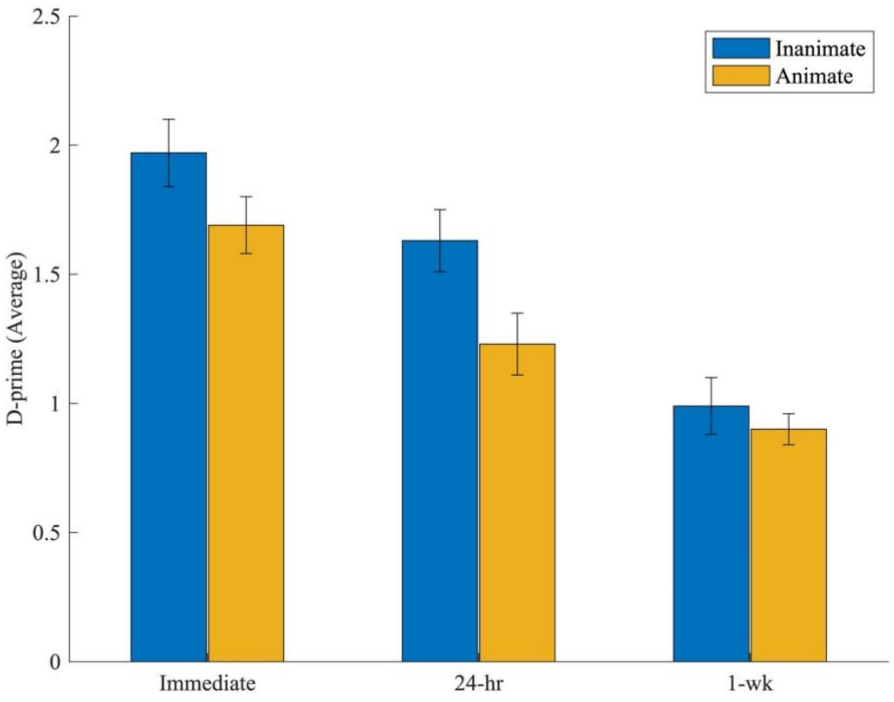

Figure 4.3 Average d' scores for target memory in the No Change condition at each test session (Immediate, 24-hour delay, 1-week delay) based on animacy (Inanimate and Animate). Error bars represent ±1 SEM.

Furthermore, when examining the response bias in the No Change condition, there was a main effect of *Time* ($F(2, 78) = 11.82$, $p < .001$, $\eta_p^2 = .23$) but not *Animacy* ($F(1, 39) = 1.59$, $p = .216$, $\eta_p^2 = .04$). Response bias significantly differed across the three test sessions. Participants adopted a slightly liberal bias at the immediate test ($M = -.03$, $SE = .05$), which was not significantly different from the liberal bias they adopted at the 24-h delayed test ($M = -.10$, $SE = .05$; $p = .298$). Differences in c were most pronounced at the 1-wk follow-up test. Participants were significantly more conservative after one week ($M = .12$, $SE = .05$) compared to the Immediate and 24-h test (all $p < .019$). Finally, we failed to observe a significant *Animacy X Time* interaction ($F(2, 78) = 1.27$, $p = .288$, $\eta_p^2 = .03$; see Figure 4.4).

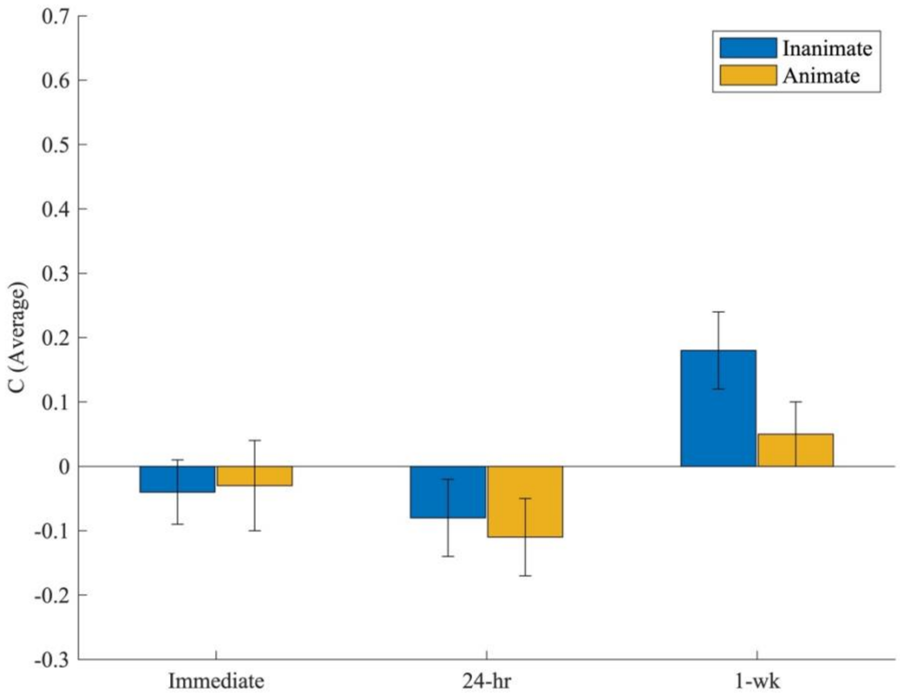

Figure 4.4 Average criterion for old/new decision in the No Change condition at each test session (Immediate, 24-hour delay, 1-week delay) based on animacy (Inanimate and Animate). Error bars represent ±1 SEM.

4.3.2.2 Contextualised Condition

We next examined the effects of animacy and time on targets that were shown in one context during study and in a new context during test (i.e., Contextualised condition; context shift decrement) based on d'. In this condition, we observed a main effect of *Time* ($F(2, 78) = 7.39$, $p = .001$, $\eta_p^2 = .16$) and Animac*y* ($F(1, 39) = 32.89$, $p < .001$, $\eta_p^2 = .46$). Like with the No Change targets, recognition linearly decreased across the week ($F(1, 39) = 11.61$, $p = .002$, $\eta_p^2 = .23$). However, when the context changed between study and test, participants were significantly better at recognising animate targets compared to inanimate targets, replicating previous animacy benefits found in the wider literature and in opposition to when the context is reinstated. For the Contextualised targets, we again failed to observe a significant *Animacy X Time* interaction ($F(2, 78) = .47$, $p = .625$, $\eta_p^2 = .01$; see Figure 4.5).

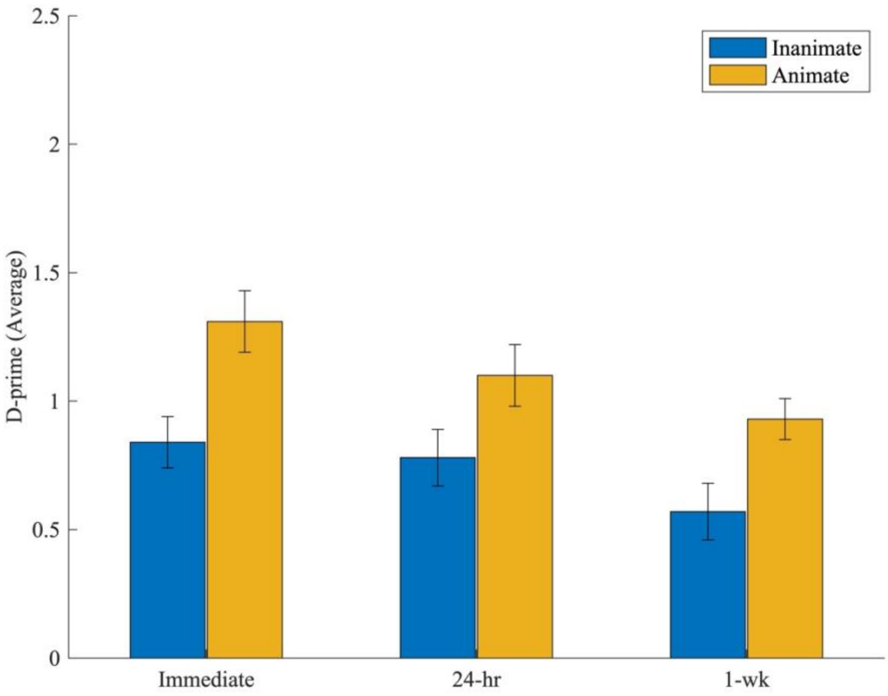

Figure 4.5 Average d' scores for target memory in the Contextualised condition at each test session (Immediate, 24-hour delay, 1-week delay) based on animacy (Inanimate and Animate). Error bars represent ±1 SEM.

Examination of the response bias for targets whose scene context changed between study and test revealed there were main effects of *Animacy* ($F(1, 39) = 49.99$, $p < .001$, $\eta_p^2 = .56$) and *Time* ($F(2, 78) = 8.46$, $p < .001$, $\eta_p^2 = .18$). Participants adopted a more conservative bias for the Inanimate targets ($M = .42$, $SE = .05$) compared to the Animate targets ($M = .05$, $SE = .05$). Response bias, averaged over both animacy conditions, was conservative at each test session, however the degree to which participants were responding conservatively changed. Participants became less conservative at the 24-h delayed test ($M = .15$, $SD = .05$) compared to the Immediate test ($M = .35$, $SE = .05$), yet they became more conservative after a 1-wk delay ($M = .22$, $SE = .05$). Whilst participants seem to be applying a liberal bias to the Animate targets at the 24-hr test, the *Animacy X Time* interaction ($F(2, 78) = .04$, $p = .964$, $\eta_p^2 < .01$) was found to not be significant (see Figure 4.6).

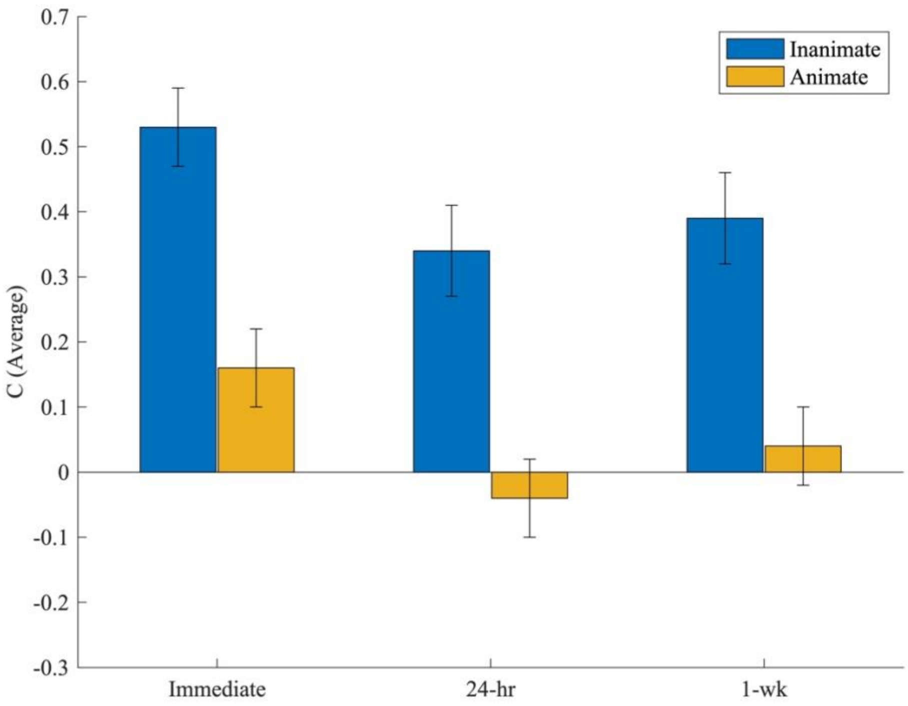

Figure 4.6 Average criterion for old/new decision in the Contextualised condition at each test session (Immediate, 24-hour delay, 1-week delay) based on animacy (Inanimate and Animate). Error bars represent ±1 SEM.

4.3.2.3 Decontextualised Condition

Finally, we examined the effects of animacy and time on targets that were consistently shown in a new scene context (i.e., Decontextualised condition), based on d’ scores. In this condition, we observed a main effect of *Time* ($F(2, 78) = 16.71$, $p < .001$, $\eta_p^2 = .30$) but not *Animacy* ($F(1, 39) = .76$, $p = .388$, $\eta_p^2 = .10$). Again, recognition linearly decreased across the week ($F(1, 39) = 26.86$, $p < .001$, $\eta_p^2 = .41$). However, the *Animacy X Time* interaction just failed to reach significance ($F(2, 78) = 3.01$, $p = .055$, $\eta_p^2 = .07$; see Figure 4.7).

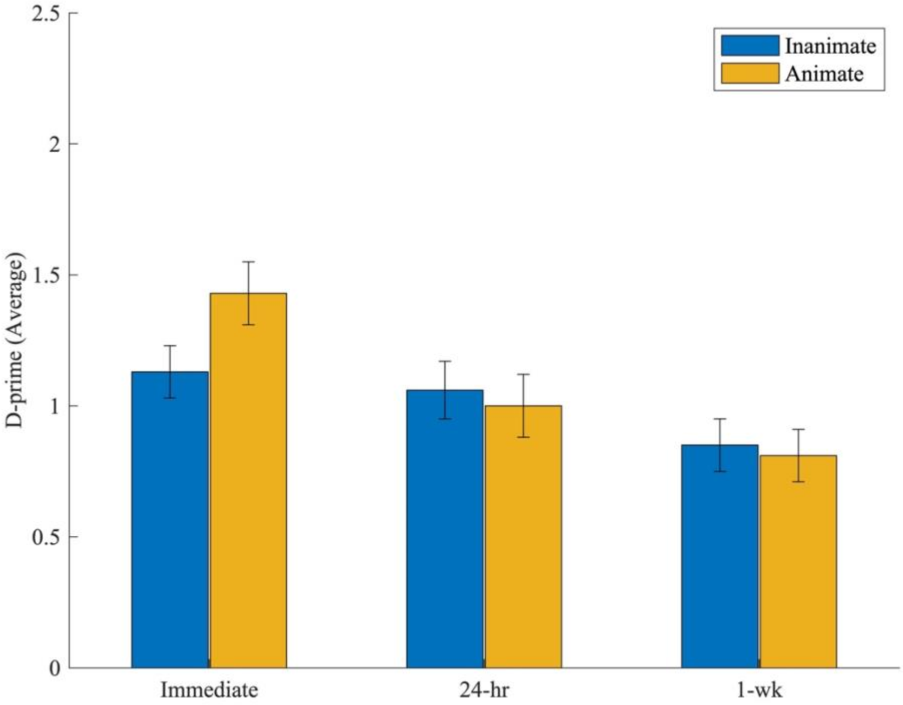

Figure 4.7 Average d' scores for target memory in the Decontextualised condition at each test session (Immediate, 24-hour delay, 1-week delay) based on animacy (Inanimate and Animate). Error bars represent ±1 SEM.

Furthermore, the results of the criterion analysis for the Decontextualised targets demonstrated there were main effects of *Animacy* ($F(1, 39) = 14.81$, $p < .001$, $\eta_p^2 = .28$) and *Time* ($F(2, 78) = 4.87$, $p = .010$, $\eta_p^2 = .11$). The criterion for the Inanimate targets ($M = .28$, $SE = .05$) was more conservative than for the Animate targets ($M = .07$, $SE = .05$). Additionally, like with the Contextualised targets, participants were less conservative at the 24-h delayed test ($M = .10$, $SE = .06$) compared to the immediate test ($M = .24$, $SE = .05$), but became more conservative again at the 1-wk test ($M = .17$, $SD = .05$). Again, the *Animacy X Time* interaction ($F(2, 78) = 1.01$, $p = .368$, $\eta_p^2 = .03$) was found to not be significant (see Figure 4.8) . A summary table for the d' and c values for the Congruent (Chapter 2) and Inanimate (current study) targets are provided below (see Table 4.3).

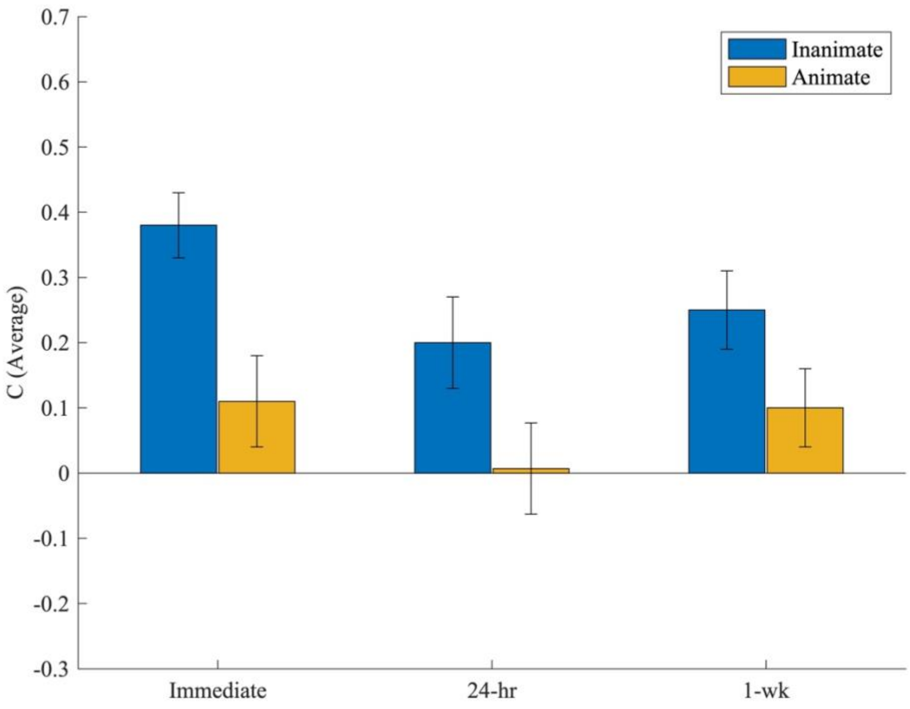

Figure 4.8 Average criterion for old/new decision in the Decontextualised condition at each test session (Immediate, 24-hour delay, 1-week delay) based on animacy (Inanimate and Animate). Error bars represent ±1 SEM.

Table 4.3 Summary table of d' and c for the Congruent targets (Chapter 2) and the same Inanimate targets (current study).

Condition	d'			c		
	Immediate	24-h	1-wk	Immediate	24-h	1-wk
Congruent						
No Change	2.09	1.55	.97	-.01	.14	.29
Contextualised	.79	.81	.47	.64	.52	.54
Decontextualised	1.01	1.00	.62	.53	.52	.47
Inanimate						
No Change	1.97	1.63	.99	-.04	-.08	.18
Contextualised	.84	.78	.57	.53	.34	.39
Decontextualised	1.13	1.06	.85	.38	.20	.25

4.3.3 Sleep Questionnaires

Descriptive statistics from the sleep questionnaires can be found in Table 4.4. Participant responses on the PSQI, the Epworth Sleepiness Scale and the Stanford Sleepiness Scale demonstrate that performance was not influenced by how tired the participants were.

Table 4.4 Average responses on the Pittsburgh Sleep Quality Index (PSQI, ranging from 1-21), Epworth Sleepiness Scale (ESS; ranging from 0-24) and the Stanford Sleepiness Scale (SSS; ranging from 1-7). Higher scores on the PSQI indicate poorer sleep quality. Higher scores on the ESS indicate excess sleepiness whilst lower scores on the SSS indicate less feelings of tiredness at that moment.

	N	Score (average)	SD
PSQI	40	5.83	± 2.67
Epworth Sleepiness Scale	40	5.21	± 3.05
Stanford Sleepiness Scale	40	2.60	± .60
SSS Immediate	40	2.33	± .80
SSS 24-hr	40	2.73	± .91
SSS 1-wk	40	2.42	± .74

4.4 Discussion

The current study expanded the investigation of context effects in visual episodic memory by further manipulating the relationship of the target to the scene. The present study presented participants with targets that were either animate or inanimate, appearing in one of three scene context conditions and tested across a week. As a reminder, the inanimate targets in the present study were identical to the congruent targets used in Chapter 2.

4.4.1 Context Reinstatement Supports Accurate Recognition

The results from the d’ scores in the No Change condition further demonstrate that the relationship between a target and its scene influences visual recognition memory. There was significantly better performance for the inanimate targets compared to the animate targets when the study context was reinstated at test, contrary to the general findings across the verbal literature that demonstrates a benefit of animate targets. For instance, Bonin et al. (2014) tested animacy effects in memory using pictures of animate/inanimate objects paired with the written word for the object. Memory was tested using free recall for the names of the pictures they had seen, and they observed an animacy advantage in their free-recall task. Whilst the Bonin et al. (2014) study is one of the very limited number of studies to examine animacy effects in visual memory, it differs substantially from the current paradigm. It appears to be the case that animacy effects are paradigm-specific, a point which will be discussed in further detail below.

Our results concerning contextual effects and inanimate targets replicate our previous findings (see Chapter 2). When the context is reinstated during encoding, tight object-scene binding occurs. If the context is reinstated at test (i.e., No Change), retrieval of the target is facilitated, consistent with the results in the Chapter 2. The present results lend further support to the argument that tight object-scene binding occurs automatically, forming strong associations between a target and contextual details when the scene context is reinstated during encoding and later reinstated during retrieval. Recognition of the animate targets, whilst not as high as the inanimate targets, was still reasonably high (i.e., Animate d': 1.28 vs Inanimate d': 1.53). It is important to note that the animate targets are congruent with the scene-context, all being plausibly found within the indoor scenes, though not typically found in a fixed location. This suggests that the animate targets are not scene specific like the inanimate targets. For example, a kettle will always be on a kitchen counter whereas a cat could be on the floor, table, or even the counters. On these grounds it appears as though animate items are less tightly bound with their scene context in memory – the scene is simply not typically a strong predictor of a particular animate target (of course though there are notable exceptions e.g., horse – stable, etc.).

4.4.2 Tight Object-scene Binding Hinders Recognition Memory

When the scene-context changed between study and test (i.e., Contextualised condition) an animate benefit emerged such that animate targets were now significantly better remembered than inanimate targets. Consistent with the findings in Chapter 2, the tight object scene binding that occurred from presenting an inanimate target in a repeated scene during encoding hindered memory for the target at test when the scene changed. The repetition of the scene-context during encoding facilitated strong object-scene binding for the inanimate targets and, when the context changed at test, memory was severely impacted since the retrieval context no longer served as a useful memory cue for the target object. More interestingly and in contrast to the No Change condition, animate targets in the Contextualised condition were significantly better remembered than the inanimate targets.

Evans and Wolfe (2022) argued that, if an object regularly moves and is present in a scene, it should be less dependent on context. For example, it is not unlikely to see a pet dog in multiple rooms in a house – they are not restricted to a single room or setting like a sofa would be.

The animate advantage observed in the Contextualised condition was arguably carried by changes in recognition scores for the inanimate targets, rather than increased performance for animate targets. Recognition scores for the animate targets was relatively unchanged, numerically speaking, when comparing the No Change (d’ = 1.28) and Contextualised (d’ = 1.11) conditions. This contrasts with a numerical drop in performance for the inanimate targets between the No Change (d’ = 1.53) and Contextualised (d’ = .73) conditions. This further supports the argument that the animate targets were not as tightly bound to their scene-contexts during encoding, making them less susceptible to interference at test when the scene changed.

4.4.3 A Lack of Consistency in Context Affects Animates and Inanimates to the Same Extent

The present study again included a Decontextualised condition in which the target object appeared in three, unique contexts (two during encoding, and a third new context at test). When the scene-context changed each time a target was encountered there were no significant differences in memory between animate and inanimate targets. Recall back to the significant congruency advantage that was observed in Chapter 2 within the Decontextualised condition. The significant increase in recognition for congruent cases, relative to incongruent cases, demonstrated that a more gist-like representation of the concept of a target was being stored in memory rather than the specific target. This meant that a semantically related scene-context at test provided a useful retrieval cue for the targets only when the target was plausible in the scene (i.e., the congruent cases). The results in the current study further support the argument that repeatedly encountering a target object in a new token of a scene type (e.g., dog in kitchen) leads to to a more gist-like representation of the target whereby a plausible scene context can still serve as a useful memory cue. It again appears to be the case that a more abstract schema of the target was being encoded rather than the specific item (e.g., a particular dog in a particular living room), meaning that the new, related scene-context at test again provides a useful retrieval cue.

4.4.4 Animacy Effects in Visual Memory

To the best of our knowledge, the present study is the first to examine animacy effects in visual episodic memory via an object-scene binding paradigm. The research that supports an animate benefit focuses on the verbal memory system or rapid visual perception. It was therefore expected that an animacy advantage would be present in all three context conditions, though this was not the case. The visual episodic memory and object-scene binding paradigm employed in the present study presents challenges when trying to make comparisons between the wider literature and the present findings. Nonetheless, the present results provide novel insights into animacy effects, particularly that the effects of animacy might be specific to the experimental paradigm employed.

The findings from the present study demonstrate that the animacy effects across the conditions were carried by performance changes in the inanimate condition. Averaging across all three test sessions, recognition scores for the animate targets were relatively stable across the three context conditions, numerically speaking (No Change d’ = 1.28, Contextualised d’ = 1.11, Decontextualised d’ = 1.08). The inanimate targets replicated the pattern of performance observed in Chapter 2 (remembering that the inanimate and incongruent stimuli were identical) with d’ averaged across test sessions fluctuating based on the context condition (No Change d’ = 1.53, Contextualised d’ = .73, Decontextualised d’ = 1.02). This provides evidence that the animate targets were not as tightly bound to their scene-contexts, supporting the argument made by Evans and Wolfe (2022). It does appear to be the case that if an object does not have a fixed location within a scene, tight object-scene binding is hindered, making the object less context-dependent at test. The animate targets were generally not affected by the changes in context, whereas recognition of the inanimate targets can be facilitated or hindered by binding to the scene-context. When scene expectations are present and useful (e.g., a kettle belongs in a kitchen), the automatic binding between an object and its context is difficult to overcome. This can have a faciliatory effect on memory (when the context is reinstated) or hinder memory (when the context changes).

The previously established animacy benefit emerged within verbal experimental paradigms and a very limited number of pictorial paradigms, yet the present findings provide evidence that animacy benefits might be paradigm specific. For example, Sun et al. (2011) and New et al. (2007) found an animacy benefit when completing a rapid categorisation task and a change detection task, respectively, using pictures as stimuli. Both studies lend credence to the idea that animates capture attention quickly, in line with the need to quickly identify threats and predators, yet this animacy benefit might be more specific to visual attention rather than long-term memory. When the task involves remembering an object long-term, as is the case in the present study, animacy has differing effects on memory which are dependent upon other factors, namely performance within the inanimate condition. Bonin et al. (2014) make the argument that animacy effects in memory paradigms persist regardless of using words or pictures, having recreated their verbal experiment with picture stimuli. The key distinction between their findings and the current findings is that their paradigm displayed the name of the animal alongside the picture and written memory was tested using free recall. The task employed in the present study was purely visual in nature and tested memory via a recognition paradigm, demonstrating a key distinction between the present study and the limited research into animacy effects in visual memory. Their study alone is not enough to generalise animacy effects in memory to the visual memory domain and the other studies using images to examine animacy effects (e.g., Sun et al., 2011; New et al., 2007) do not test memory for the targets.

Only one recently published study has examined animacy effects within a context-binding paradigm, though it is important to note that they used word-pair associations rather than pictures for their stimuli (Schreiner et al., 2023). Schreiner et al. (2023) employed an open-/closed-loop binding task in which participants had to encode meaningful associations between three associative words. Within the open-loop condition, participants were presented with two overlapping pairs of words, sharing a common associate whereas in the closed-loop condition all three pairwise combinations were presented. Interestingly and in line with the present findings, they observed that an animate benefit emerged in some cases but was absent in other cases. For example, performance was lower for animal-location pairs compared to object-location pairs. Even when using animacy as a between-subjects factor (Experiment 3), they observed “strong” evidence against animacy effects, with performance being similar between the animate/inanimate conditions. They argue that the nature of asking participants to imagine meaningful interactions between the pairs may have inhibited the emergence of an animacy effect within their paradigm due to the task requirements boosting performance in the inanimate condition. This indeed seems to be the case in the present study whereby the task requirements inhibited the expected animacy effects in some of the context conditions, due to changes in recognition for the inanimate targets. Reinstatement of a consistent scene context facilitated tight object-scene binding for the inanimate targets, facilitating memory whereas the same tight binding hindered memory when the context changed between study and test. The animate targets were not as tightly bound to their contexts, making them relatively unaffected by whether the context was reinstated or changed at test. This led to an inanimate advantage in one case (i.e., No Change) and an animate benefit in the other (i.e., Contextualised). To say that animacy effects were driving this pattern of results is not entirely accurate as the large hinderance to performance for the inanimate targets drove this effect. This provides support for Schreiner et al.'s (2023) argument and highlights the importance of acknowledging task requirements when discussing how specific effects operate.

4.5 Conclusion

The present experiment extended the findings of the experiment reported in Chapter 2 by examining how animacy and context impacts recognition for elements of visual episodic events. These results taken together with the results from the previous study suggest that the relationship of the target object to the scene influences the extent to which context plays a role in retrieval. When it comes to the influence of context on animate targets, the pattern of results differs from those observed for the incongruent targets in Chapter 2. This is likely due to the fact that animate objects (e.g., dogs, cats, birds) are more likely to be encountered in a variety of contexts in real-world scenarios making them less context-dependent. However, the present results further support the findings from Chapter 2 that the strength of object-scene binding contributes to memory for specific elements of the episode. The tight object-scene binding that results from repeating a scene context during encoding (and further reinstating it during test) facilitates memory for the targets, though it is more beneficial for inanimate targets. Whilst tight binding facilitates memory when the context is reinstated, it hinders memory when a change in context occurs between study and test. However, the extent to which this hinders memory is determined by the relationship of the target to the scene. In this case, animate targets are better remembered than inanimate targets most likely due to them regularly being encountered in a variety of scenes. Finally, further support is gained for the idea that presenting a target in different scene contexts of the same type leads to a more gist-like or schematic representation of the target object. This was evidenced by no significant difference between animate and inanimate targets in the Decontextualised condition. For both types of targets, the new context at test matches the gist representation of the targets, facilitating retrieval. One can therefore argue that elements in a visual episode (e.g., an object in a scene) are not independent from the context they appear in which can both hinder and support memory for the specific elements.

Appendix

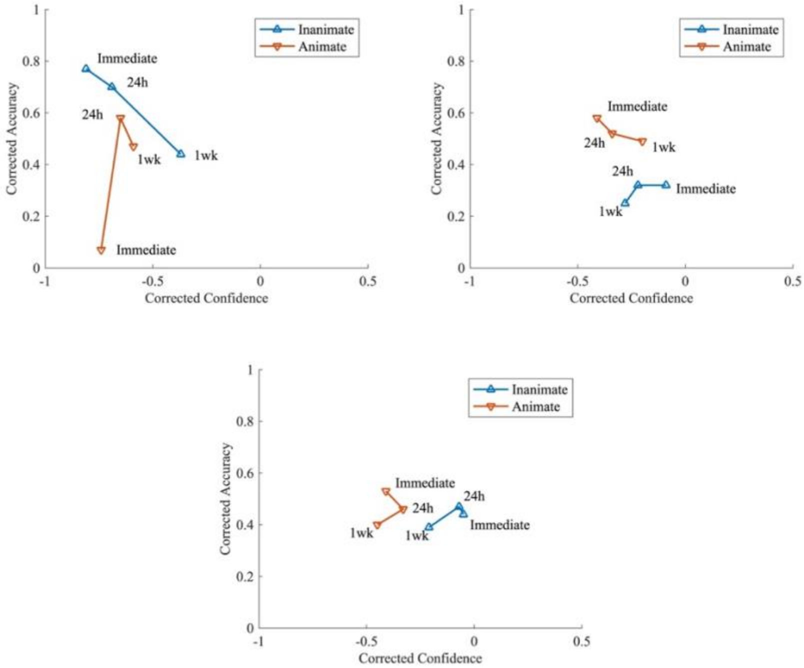

Figure 4.9 Average corrected accuracy and corrected confidence for Congruent and Incongruent target words in sentences at each of the three test sessions for the No Change (panel a), Contextualised (panel b) and Decontextualised condition (panel c).

5. Neural Mechanisms of Episodic Memory

5.1 Introduction

Based on the work presented in the previous chapters, it is clear that context plays an important role in visual episodic memory. However, the extent to which context impacts our ability to accurately recognise previously encountered targets depends on the relationship between the object and the background scene. The question that arises from these behavioural results pertains to what neural mechanisms support the formation, consolidation, and retrieval of visual episodic memories. Additionally, it is unclear whether the results obtained in the previous chapters arise from neural processes at encoding, retrieval or both.

5.1.1 Advantages of EEG Methods and Time Frequency Analysis

The previous chapters shed light behaviourally on people's ability to accurately store and retrieve elements of episodic memories. However, another way that helps to elucidate the mechanisms underlying this effect is to use neuroscientific approaches to measure brain activity whilst participants encode and retrieve the targets. Using electroencephalogram (EEG), electrical activity in the brain can be measured whilst participants complete the behavioural tasks, and the pattern of activity can be analysed, allowing direct measurements of neurocognitive processes at the neuronal level (Cohen, 2014). Early EEG research employed two main techniques when examining changes in neural data: event related potentials (ERPs) and Fourier transform analyses. ERP analyses measure fluctuations in brain activity voltage millisecond by millisecond (Woodman, 2010), allowing for highly accurate temporal analyses that require relatively little data processing (Cohen, 2014). It is primarily focused on 'components' of a more complex waveform that varies by polarity, timing, distribution and sensitivity to the task (Woodman, 2010). For example, the N400 component occurs consistently ~200-500 ms after stimulus onset in response to linguistic characteristics like word frequency, congruency and repetition, and is characterised by a decrease in the amplitude of the waveform or a 'negative deflection' (e.g., Besson et al., 1992; DeLong et al., 2005; DeLong & Kutas, 2020; Dini et al., 2022; Hamm et al., 2002). Whilst ERP components have been well-established, they come with limitations in understanding cognitive functions. For example, ERPs provide little information regarding the dynamic nature of the EEG data, particularly when it comes to task-related information in the signal, making null results difficult to interpret (Cohen, 2014).

An alternative method of analysis involves Fourier transform of the EEG signals. Fourier transform analyses, on the other hand, are able to capture information about oscillations in brain activity, measuring the amount of activity present within a given frequency band, commonly referred to as 'power' (i.e., amplitude squared; Morales & Bowers, 2022). The main limitation of Fourier transform analyses is that it assumes the EEG signal is stationary, meaning temporal information within the EEG signal becomes lost (Morales & Bowers, 2022). This becomes problematic when looking at neuronal changes across time (i.e., experimental trials lasting more than few hundred milliseconds) since psychophysiological processes fluctuate over time (Cohen, 2014; Morales & Bowers, 2022). However, recent advances in EEG analyses have found ways of overcoming these limitations in order to capture a more accurate snapshot of neural processes during experimental tasks (i.e., time frequency analysis).

One such method of analysis, time frequency analysis, involves separating neural activity into pre-defined frequency bands: delta (1-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz) and gamma (30+ Hz; Iber et al, 2007). Time frequency analysis proves advantageous over ERP and Fourier transform analyses because it allows one to examine EEG data as a multidimensional signal that captures task-relevant dynamics in the data (Cohen, 2014). Moreover, it allows for the examination of temporal dynamics present in the EEG oscillatory activity, namely frequency, power and phase (Morales & Bowers, 2022). This allows for a more fine-tuned examination of the "dynamic" changes in activity with more temporal precision rather than examining brain activity as voltage spikes at a specific point in time (i.e., ERPs; Morales & Bowers, 2022). One final benefit to the time frequency approach is that it allows statistical analyses similar to those used on behavioural data in order to understand what neural activity drives the behavioural effects. The main focus of the current research pertains to what role theta activity plays whilst encoding and retrieving episodic memories. The rationale behind this particular frequency band is explained below.

5.1.2 Theta Activity (4-8 Hz) in Episodic Memory

What role different frequency bands play in memory has long been a topic of study, with individual frequency bands being shown to support differing aspects of memory. For example, beta activity has been associated with memory reactivation (Zhu et al., 2019) and increases in alpha activity have been shown to reflect the inhibition of irrelevant task information (Hanslmayr et al., 2012). The present study specifically focuses on theta activity (4-8 Hz) since previous studies have demonstrated that increases in theta activity may play a key role in episodic memory formation (Atienza et al., 2011; Friese et al., 2013; Griffiths et al., 2019; Hanslmayr et al., 2012; Roux et al., 2022), item-context binding (Hanslmayr et al., 2009, 2016; Staudigl & Hanslmayr, 2013), and in associative binding (Köster et al., 2018). For example, Köster et al. (2018) presented participants with a series of words/pictures superimposed on a coloured square and asked them to form an association between the target and the colour. During retrieval, targets and foils were presented on a grey background and participants were asked to respond with whether the target was old or new. If the target was old, they were asked to indicate whether or not they were able to retrieve the associated colour. Memory performance and theta activity were higher for pictures compared to words. Moreover, they observed an increase in theta activity for subsequently remembered targets, particularly over frontal electrode sites. The increase in theta activity was accompanied by a selective increase in theta-gamma coupling when the pictures were recalled with a colour association in comparison to when no colour was retrieved or to words. They argue that the increase in theta activity and theta-gamma coupling reflected the successful binding of visually rich information since no such pattern of neuronal activity was observed for the verbal stimuli.

Similarly, Staudigl and Hanslmayr (2013) presented participants with a series of words superimposed on short movie clips and had participants complete an old/new task for the word/video pairs at test. During the testing phase, some of the target words appeared on the same clip they were presented on during the study phase whilst some of the target words appeared on new clips. They found that higher theta power during encoding was an accurate predictor of successful retrieval only when the study and test contexts matched, and conversely that theta power increases during encoding were detrimental to memory when there was a change in context at test. It seems reasonable to assume that, within our context-binding visual paradigm, changes in theta power should be observed in response to context reinstatement (i.e., No Change condition) and context changes (i.e., Contextualised and Decontextualised condition). What remains unclear from previous studies is whether these effects will be present during encoding when images are presented multiple times, or whether these effects will only be present during retrieval, as is the case in previous work where images are presented once during study and once during test.

Changes in theta power have also been observed when comparing neural activity for items that are semantically congruent to items that are semantically incongruent, though a majority of the research in this area focuses on verbal material. In a review by Maguire and Abel (2013), focusing on language development studies, theta activity appears to be linked to semantic integration, highlighting that theta is particularly sensitive to congruency effects. Research suggests that theta activity increases when presented with incongruent information (Dini et al., 2022; Schneider & Maguire, 2018) and in response to unexpected stimuli (Klimesch et al., 1994). For example, Schneider and Maguire (2018) had participants listen to 160 sentences which were either grammatically correct or contained a semantic error. Participants were asked whether the action-agent pairing was “plausible” or “not plausible” and whether the pairing was correct/incorrect. They found that theta activity significantly increased when the sentences contained a semantic error (i.e., incongruent sentence), reflective of increased demands on the memory system when trying to integrate semantic violations into coherent memory traces. However, some studies have shown that congruent targets can also elicit stronger theta responses relative to incongruent targets which leads to faster reaction times at retrieval (Atienza et al., 2011). In their study, participants were presented with a grid that contained a face (either famous or non-famous) and a congruent/incongruent fact about the person (random biographical information was provided for non-famous faces). Memory for the faces was tested using a visual-spatial test regarding whether the face was in the original position or in a new position. They found that semantically congruent face-location associations elicited stronger theta activity during encoding and were correlated with faster reaction times during retrieval. Atienza and colleagues (2011) suggest that theta and congruency may mediate interactions between brain regions during successful encoding by redirecting attention towards task/goal-relevant information. Due to the conflicting nature of theta effects in congruency studies, the present study additionally sought to examine whether theta power changes would be present for congruent versus incongruent targets during encoding and retrieval.

5.1.3 The Role of Sleep in Memory

In addition to theta's role in memory formation and retrieval of episodic memories, theta is believed to play a role in hippocampal (HC) to neocortical (NC) communication (Fogel et al., 2009; Reiner et al., 2014), a crucial step in memory consolidation. The process of consolidation involves stabilising fragile memory traces into more durable representations within the long-term memory networks, usually during quiet periods where no new information is being encoded (Sutherland, 2000). Sleep has consistently been shown to facilitate memory consolidation by stabilising newly encoded information, providing an optimal offline environment free of new, competing information. Theoretical accounts have established that sleep plays a beneficial role in memory consolidation, though there is debate as to whether sleep plays an active or passive role in the process (e.g., simply defending against retrograde interference; Lahl et al., 2008). One widely accepted theory (the active systems consolidation hypothesis) suggests an active role of sleep in consolidation whereby memory traces are reactivated during slow wave sleep (SWS) leading to the transfer of the memory traces from the HC to to the NC which, in turn, strengthens the cortical connections (Born & Wilhelm, 2012; Dudai, 2004; Klinzing et al., 2019; MacDonald & Cote, 2021; Zhang, 2022). Conversely, the synaptic homeostasis hypothesis argues that sleep plays a more passive role, with consolidation being facilitated when synaptic downscaling returns the synaptic weight to a 'baseline' level, resetting the brain for new learning the next day (Tononi & Cirelli, 2003, 2006). Whilst these two key theories offer opposing views to the role sleep plays in consolidation, they both support the idea that preventing interference of new information is a key element involved in consolidation.

In the studies conducted and presented in Chapters 2-4, the learning and immediate test phases occurred at different times of day for each participant, and sleep did not occur immediately after learning. Across the studies presented in the present thesis, performance decreased as the time between study and test increased, suggesting overnight sleep did not influence memory for the targets. Graded retroactive interference theories suggest that forgetting at test will be most severe if new information is presented shortly after encoding (Yonelinas et al., 2019). This is due to the newly presented information having greater contextual overlap (i.e., temporal context). This has been found to occur regardless of how similar the new information is to the study material (Yonelinas et al., 2019). Moreover, the contextual binding model of episodic memory posits that sleep's role might simply be to protect the memory traces from further interference incurred during wakefulness (MacDonald & Cote, 2021; Yonelinas et al., 2019). The contextual binding model further suggests that episodic memory is dependent on the HC, with the HC playing a key role in binding the item and its context into a cohesive memory trace (Yonelinas et al., 2019). If the contextual binding model is correct, then it is possible that there was no benefit of sleep at the 24-hour test in any of the studies presented in Chapters 2-4 due to interference from incoming information the participants encountered between the experimental session and their normal bedtime. The present study addressed this by having participants come into the lab to complete the experiment. Participants completed a study phase followed by an immediate test, identical to the procedure in the previous three studies. However, instead of taking the second test after a 24-hr delay, participants spent a 2-hr delay period in the lab. The second test occurred after the 2-hr delay, eliminating the time-of-day differences across participants.

5.1.4 Present Research Aims

The interesting question that remains from the previous chapters is whether the effects of context and congruency are supported by neuronal processes at encoding, retrieval or both. In Chapter 2, participants encoded a series of target objects embedded in real-world, indoor scenes. The basic design was that used in Experiment 1 (i.e., Chapter 2) in which congruent and incongruent targets were presented in the same No Change, Contextualised and Decontextualised conditions as before.

It is clear from the behavioural data in Chapter 2 that, at least for the Decontextualised condition, processes at retrieval facilitate whether the target will be correctly recognised; though what specific processes facilitate this remain unclear. Context-binding does not take place during encoding within the Decontextualised condition due to the scene-context changing each time the target is encountered. For congruent targets, this leads to the gist of the abstract description of a target (a schema) being stored in memory rather than the specific item (e.g., sofa in living room as opposed to specific sofa in specific living room). This gist description then acts as a useful retrieval cue at test. For incongruent targets though no such gist representation is ever generated because by their very nature incongruent targets violate scene expectations. On these grounds there is no useful gist description to act as a retrieval cue. In addition, it is also possible that the poor performance with the incongruent targets in the Decontextualised condition was due to disruption during encoding. It is this possibility that will be examined in the current experiment.

More particularly, EEG recordings were obtained during encoding and retrieval, using the same stimuli and a similar experimental paradigm to those presented in Chapter 2, in order to examine neural activity during the two main experimental phases. This was used to determine whether processes during encoding and/or retrieval underlie the congruency effects observed in the previous chapter.

5.2 Methods

5.2.1 Participants

We collected data from 50 healthy participants aged 18-30 (32 female). Two participants were removed due to technical issues with MATLAB during the experiment, one was excluded for having participated in a sleep deprivation study the night before, and two were excluded for sleeping less than 1-hr during the 90-min nap opportunity. This left us with a final sample size of 45 participants. All participants were fluent English speakers, had normal or corrected-to-normal vision, no special visual characteristics (e.g., colour blindness or astigmatism), and had no history of drug or alcohol abuse. They had no history of hearing problems, dyslexia, or neurological, psychiatric or sleep disorders, and they did not have a job that involved alternating shift patterns (i.e., day and night shifts). All participants provided written informed consent and filled in demographic questionnaires before the study. The study was approved by the Research Ethics Committees of the Department of Psychology, University of York. Participants received either 4 course credits or a £30 Amazon voucher as compensation upon completing the study.

5.2.2 Materials

The stimuli used were the same stimuli used in the previous behavioural study on context and congruency effects in visual episodic memory (see Chapter 2).

5.2.3 Experimental Design

The current study was conducted using a mixed design. The within-subjects factors were identical to Chapter 2. The three main within-subjects factors examined were Context (No Change, Contextualised, Decontextualised), Congruency (Congruent, Incongruent) and Time (Immediate, Delayed). The between-subjects factor was Sleep vs Wake, determined by what the participants were asked to do during the 2-hr period between the immediate and delayed test.

5.2.4 Equipment

5.2.4.1 Experimental Tasks

All tasks were conducted in a Windows PC and participant responses were recorded using the left and right arrow keys. The experiment was implemented in MATLAB 2019a (The MathWorks, Inc) using Psychtoolbox 3.0.15 (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997).

5.2.4.2 Electroencephalography

EEG recordings were attained using two Embla N7000 systems with REMLogic 3.4 software. Gold-plated electrodes were attached to the scalp according to the international 10-20 system at frontal (F3 and F4), central (C3 and C4), parietal (P3 and P4), temporal (T3 and T4) and occipital (O1 and O2) locations and were referenced to the linked mastoids. Left and right electrooculography electrodes were attached, as were electromyography electrodes at the mentalis and submentalis bilaterally, and a ground electrode was attached to the forehead. All electrodes had a connection impedance of < 5 kΩ immediately before any EEG data were collected (i.e., before the study phase, first test phase, nap and second test phase). Any electrodes that were above this threshold were replaced and rechecked. All online signals were digitally sampled at 200 Hz.

5.2.5 Procedure

5.2.5.1 Experimental Timeline

Participants were pseudo-randomly assigned to either the Wake or Nap condition. Participants arrived at the lab ~12.00. Upon arrival, they were asked to fill in the Pittsburgh Sleep Quality Index (Buysse et al., 1989) and the Epworth Sleepiness Scale (Johns, 1991). Participants were then wired-up for EEG recordings. See Figure 5.1A for the full timeline.

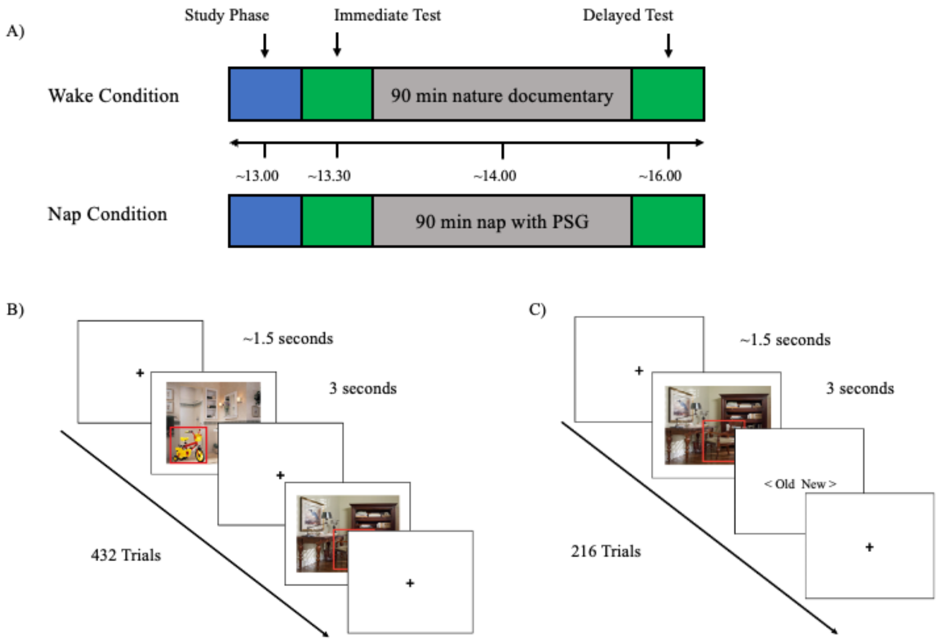

Figure 5.1 Experimental timeline and procedure. A) Experimental timeline. Participants arrived at the lab and were wired up for EEG data collection around 12.00. The experiment started immediately after wiring up had finished. B) Study phase procedure. Participants saw 432 images for 3s separated by a ~1.5s fixation cross. C) Test phase procedure. Participants were shown an image for 3s and then were prompted to make an old/new judgement. There were 216 trials per test phase.

5.2.5.2 Study Phase

The study phase started ~13.00. Participants were shown 432 images (216 targets each presented twice) for 3s each. Each image was followed by a ~1.5s fixation cross. The timing of the fixation cross was jittered (i.e., .1-.5 ms was added to a minimum 1.5 seconds) to ensure neural activity did not synchronise with the image presentation due to predictability of stimulus onset. The congruency and context conditions for the targets were the same as those used in the previous behavioural study (see Chapter 2). Participants were asked to remain as still as possible whilst the image was on the screen and to memorise the object appearing in the red bounding box. See Figure 5.1B for the study phase procedure.

5.2.5.3 Test Phase

The immediate test phase started ~13.30 and the delayed test phase started ~16.00 (after the 90-min period of quiet wakefulness or nap). Participants were shown an image for 3 seconds before being prompted to respond whether the target object in the red bounding box was old or new (see Figure 1C). A jittered fixation cross appeared between each image for ~1.5 seconds. Half of the targets were tested at the immediate test and half were tested at the delayed test. Each target was only tested once (i.e., at either the immediate or delayed test phase). At each test, there were 108 targets and 108 foils for a total of 216 trials per test phase. Participants were asked to remain as still as possible whilst the image was on the screen and to make their response using the left and right arrow keys after the image had disappeared.

5.2.6 Data Analysis

5.2.6.1 Behavioural Data

The same approach to data analysis adopted in Chapters 2-4 were adopted here. Statistical comparisons were made between the congruent and incongruent stimuli separately for each of the three context conditions. Statistical comparisons between the three context conditions could not be assessed due to the same stimuli from Chapter 2 being used in the present study.

5.2.6.2 EEG Data

5.2.5.1.1 Preprocessing.

EEG data were preprocessed and analysed in MATLAB 2022b (The MathWorks, Inc) using the Fieldtrip toolbox (Oostenveld et al., 2011). All 10 channels (F3, F4, C3, C4, P3, P4, O1, O2, T3 and T4) were included in our analyses. Data were highpass filtered (0.3 Hz) and re-referenced to linked mastoids. Data were segmented into trials from -1.5 s to 3 s around stimulus onset. Eye-blinks and cardiac components were identified and removed from scalp electrodes using an Independent Component Analysis. Any noisy channels were then interpolated via a weighted-average of their nearest neighbours.

5.2.6.1.2 Time Frequency Analyses.

Time-frequency representations (TFRs) were calculated across all frequencies using a lowpass filter of 30 Hz. Data were then convolved with a 5-cycle Hanning taper in 0.5 Hz frequency steps and in 5 ms time steps using an adaptive window. The adaptive window length decreased with increasing frequency. Artefact rejection was done using a data-driven approach whereby power values for each trial that exceeded the 85th percentile across all time/frequency points were removed from each participant's dataset. TFRs were then converted into percent power change relative to a -400 to -200 ms pre-stimulus baseline window for each frequency band. The baseline window was chosen to decrease the likelihood of contamination from the baseline period into the post-stimulus activity with particular regard to lower frequencies (i.e., theta; 4-8 Hz).

5.2.6.1.3 Statistics.

Nonparametric statistical analyses were conducted on the TFRs using dependent-samples t-tests, corrected for multiple comparisons using a cluster permutation method (Maris & Oostenveld, 2007). Clusters were defined by channel and time, averaging across the frequency band of interest (i.e., theta; 4-8 Hz). Note, it is typically the case that the analysis identifies clusters in a singular hemisphere rather than cross-hemisphere, though differences based on localisation are beyond the scope of the current thesis. The analyses used 1000 randomisations, and the cluster threshold was set to $p < .05$. A factorial approach was used to examine the effects of presentation order (encoding), test order (retrieval), congruency, and context on the neural correlates of encoding and retrieval. First, grand averages of the TFR for the different conditions of interest were calculated for encoding and retrieval. These contrasts were then entered into the cluster-based permutation analyses.

5.3 Results

5.3.1 Behavioural Data

5.3.1.1 Hits and False Alarms

The hit rate (HR) and false alarm rate (FAR) at the immediate test were z-transformed and plotted, consistent with Chapters 2-4 (see Figure 5.2). The pattern of results replicates the findings in Chapter 2, using the same stimuli (see Table 5.1), providing robust evidence that context can both hinder and facilitate episodic memory retrieval. Performance is again numerically best (regardless of whether the target is semantically congruent or semantically incongruent with the scene context) when the scene-context is reinstated between study and test. Changing the context between study/test leads to a steep decline in performance aiding further evidence to the classic context reinstatement and context shift decrement effects. If the target is consistently encountered in a new scene context (Decontextualised condition), performance relative to the Contextualised condition is slightly better for the Congruent targets but is further hindered for the Incongruent targets.

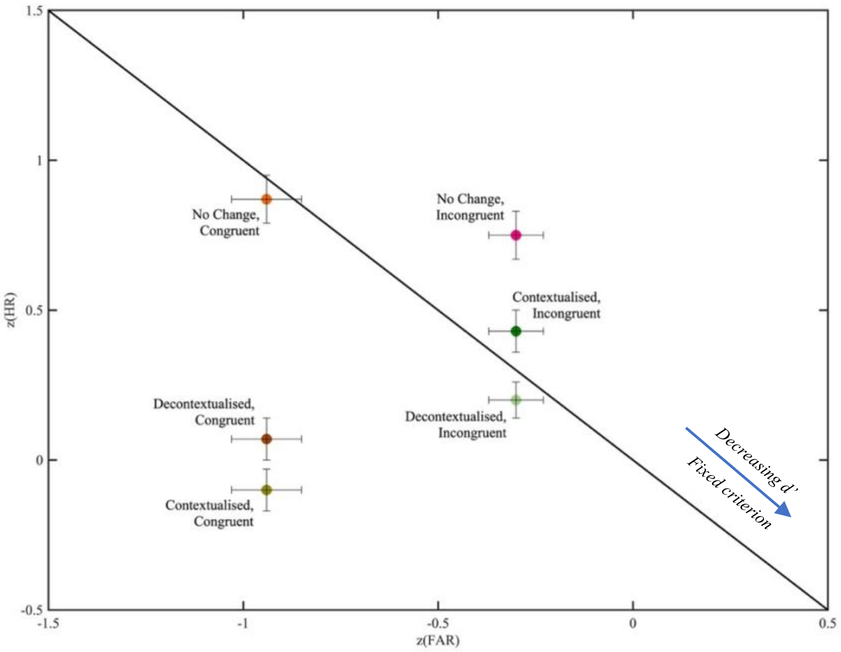

Figure 5.2 Memory performance at the Immediate test represented by z-transformed Hit Rate plotted against z-transformed False Alarm Rate. Better performance is indicated by proximity to the upper left corner and the diagonal black line indicates a neutral criterion. Each data point represents the average for one condition in the experiment. Error bars represent ±1 *SEM*.

Table 5.1 Average d' scores for each condition in the present experiment and from the corresponding experiment in Chapter 2.

Condition	Current Study (d')	Chapter 2 (d')
Congruent		
No Change	1.82	2.09
Contextualised	.84	.79
Decontextualised	1.01	1.01
Incongruent		
No Change	1.05	1.21
Contextualised	.75	.82
Decontextualised	.50	.52

Note. The stimuli used in the present study are identical to the stimuli used in Chapter 2. The study and test procedure were identical. The only changes to the experiment being the experimental timeline to accommodate the addition of the nap vs wake delay groups, which resulted in more trials at test (i.e., memory was tested across two test sessions rather than three).

5.3.1.2 No Change Condition

The first behavioural results examined the effects of congruency and time on targets that were consistently shown in the same scene context (i.e., No Change condition; context reinstatement) via d'. Main effects of *Time* ($F(1, 43) = 24.30$, $p < .001$, $\eta_p^2 = .36$) and *Congruency* ($F(1, 43) = 127.04$, $p < .001$, $\eta_p^2 = .75$) were observed. Recognition linearly decreased across the two test sessions and Congruent targets were significantly better remembered than Incongruent targets (See Figure 5.3). Interestingly, there was no significant *Congruency X Time* interaction ($F(1, 43) = .50$, $p = .502$, $\eta_p^2 = .01$) and there was no significant difference between the *Nap* and *Wake* groups ($F(1, 43) = .23$, $p = .634$, $\eta_p^2 < .01$).

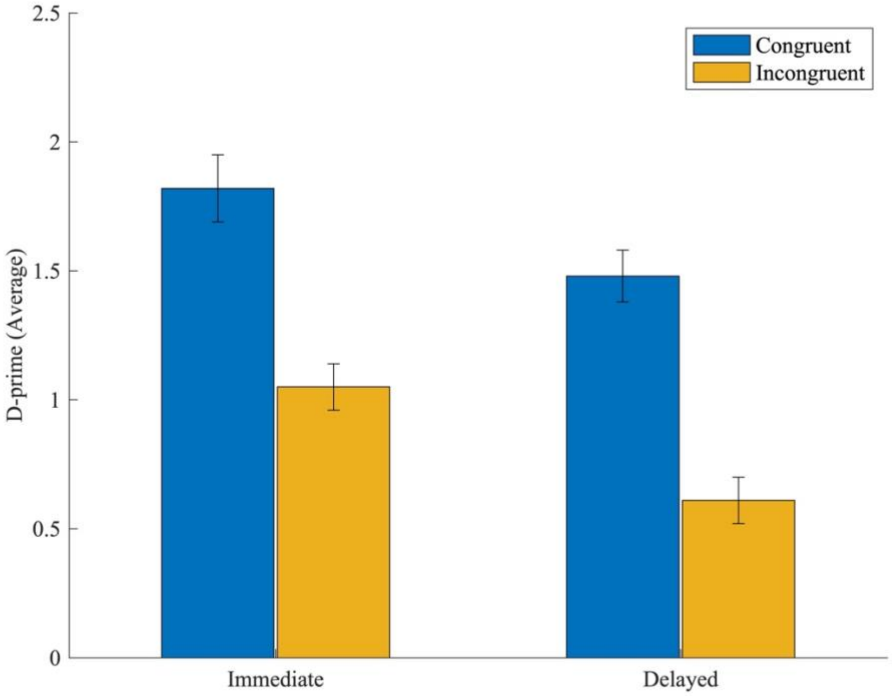

Figure 5.3 Average d' scores for target memory in the No Change condition at each test session (Immediate and Delayed) based on congruency (Congruent and Incongruent). Error bars represent ±1 SEM.

Response bias in the No Change condition was next examined. Again, there were main effects of *Congruency* ($F(1, 43) = 12.68$, $p = .001$, $\eta_p{}^2 = .23$) and *Time* ($F(1, 43) = 45.91$, $p < .001$, $\eta_p{}^2 = .52$), consistent with findings from Chapter 2 (see Figure 5.4). Participants tended to adopt a more conservative criterion threshold for the Congruent targets ($M = .15$, $SE = .06$) when compared with the Incongruent targets ($M = -.06$, $SE = .06$). Participants were liberally responding at the Immediate test ($M = -.09$, $SE = .05$), but applied a more conservative criterion after a 2-h delay ($M = .19$, $SE = .06$). Finally, there was no significant *Congruency X Time* interaction ($F(1, 43) = 1.27$, $p = .266$, $\eta_p{}^2 = .03$) and no significant difference in c between the *Nap* and *Wake* groups ($F(1, 43) = 1.46$, $p = .233$, $\eta_p{}^2 = .03$).

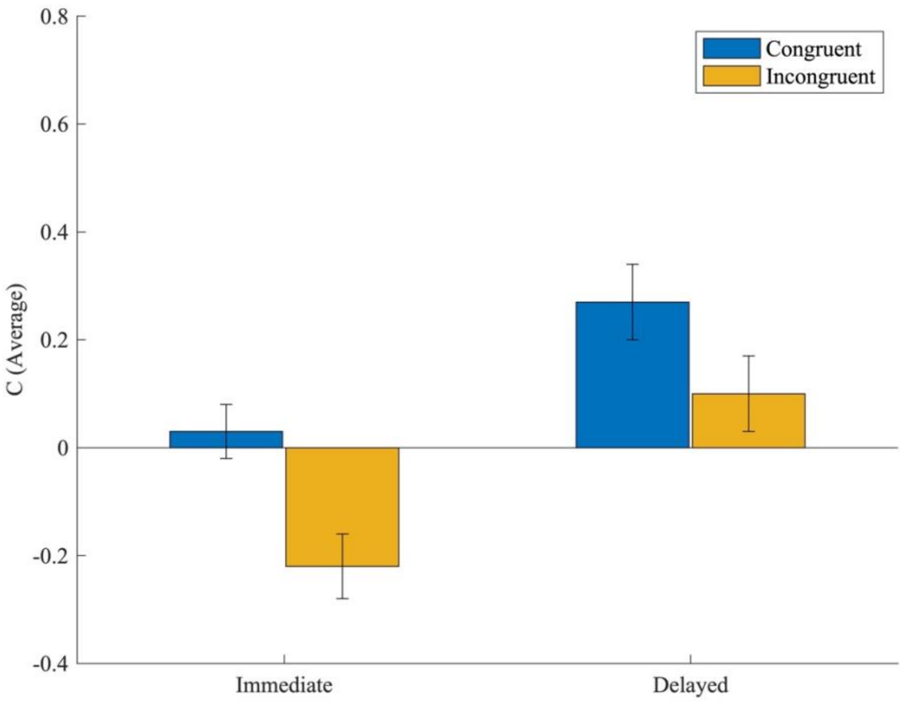

Figure 5.4 Average criterion for old/new decision in the No Change condition at each test session (Immediate and Delayed) based on congruency (Congruent and Incongruent). Error bars represent ±1 SEM.

5.3.1.3 Contextualised Condition

The effects of congruency and time on targets that were shown in one context during study and in a new context during test (i.e., Contextualised condition; context shift decrement) based on d’ were examined next. Again replicating the findings in Chapter 2, a main effect of *Time* ($F(1, 43) = 6.02$, $p = .018$, $\eta_p^2 = .12$) was observed, with a linear decrease across the 2-h delay (see Figure 5.5). However, the main effect of *Congruency* just failed to reach significance ($F(1, 43) = 3.52$, $p = .067$, $\eta_p^2 = .08$). Similar to the No Change targets, there was no significant *Congruency X Time* interaction ($F(1, 43) = .04$, $p = .848$, $\eta_p^2 < .01$) and no significant difference between the *Nap* and *Wake* group($F(1, 43) = .01$, $p = .480$, $\eta_p^2 = .01$).

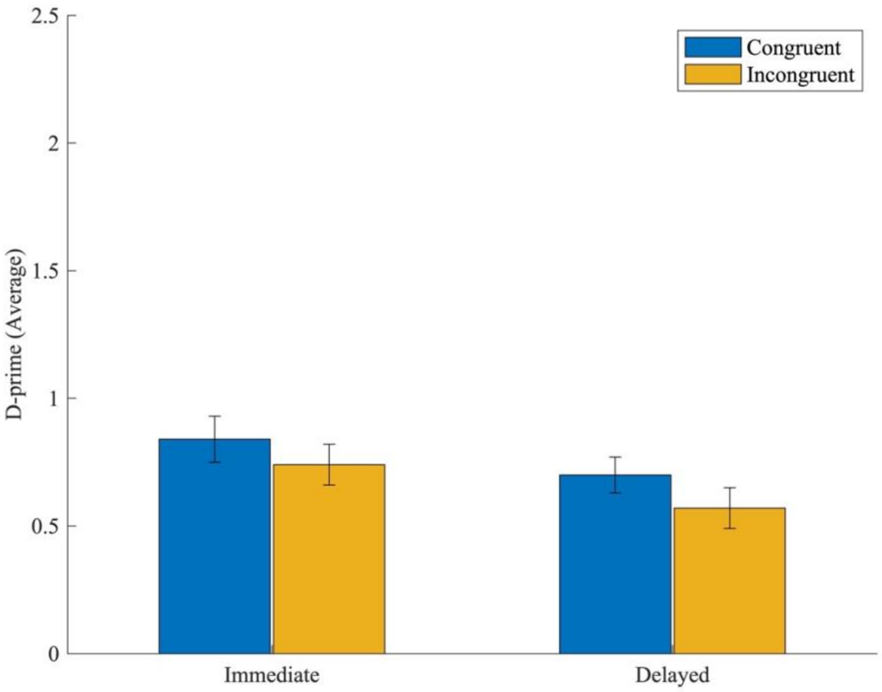

Figure 5.5 Average d' scores for target memory in the Contextualised condition at each test session (Immediate and Delayed) based on congruency (Congruent and Incongruent). Error bars represent ±1 SEM.

Examination of the criterion threshold for targets whose scene context changed between study and test revealed there were main effects of both *Congruency* ($F(1, 43) = 95.62$, $p < .001$, $\eta_p^2 = .69$) and *Time* ($F(1, 43) = 15.90$, $p < .001$, $\eta_p^2 = .27$). Much like No Change targets, participants adopted a conservative criterion threshold for the Congruent targets ($M = .59$, $SE = .06$) whilst they applied a liberal criterion for the Incongruent targets ($M = -.02$, $SE = .06$). Additionally, participants became more conservative as the length of time between study and test increased (Immediate: $M = .23$, $SE = .06$, Delayed: $M = .39$, $SE = .06$). Furthermore, the *Congruency X Time* interaction ($F(1, 43) = .39$, $p = .536$, $\eta_p^2 = .01$) was found to not be significant (see Figure 5.6) and there were no significant differences between the *Nap* and *Wake* group ($F(1, 43) = 1.18$, $p = .284$, $\eta_p^2 = .03$).

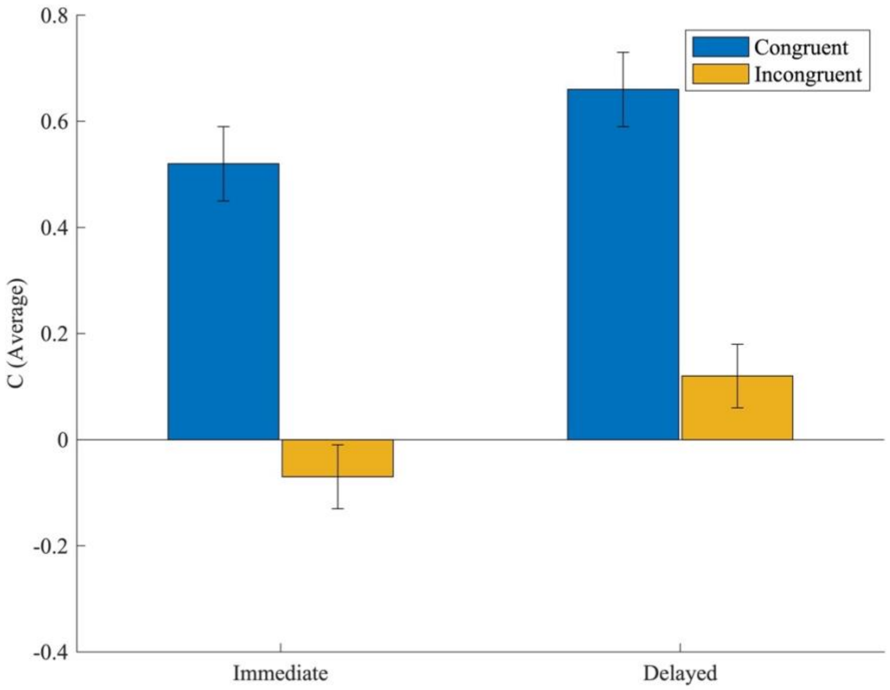

Figure 5.6 Average criterion for old/new decision in the Contextualised condition at each test session (Immediate and Delayed) based on congruency (Congruent and Incongruent). Error bars represent ±1 SEM.

5.3.1.4 Decontextualised Condition

Finally, the effects of congruency and time on targets that were consistently shown in a new scene context (i.e., Decontextualised condition), based on d' scores, was examined. In this condition, a main effect of *Congruency* ($F(1, 43 = 75.59$, $p < .001$, $\eta_p^2 = .64$) was observed, with memory performance being higher for Congruent targets. However, there was no significant main effect of *Time* ($F(1, 43) = 1.52$, $p = .22$, $\eta_p^2 = .03$). Furthermore, there was no significant *Congruency X Time* interaction ($F(1, 43) = .94$, $p = .339$, $\eta_p^2 = .02$; see Figure 5.7) and no significant difference between the *Nap* and *Wake* groups ($F(1, 43) = .13$, $p = .721$, $\eta_p^2 < .01$).

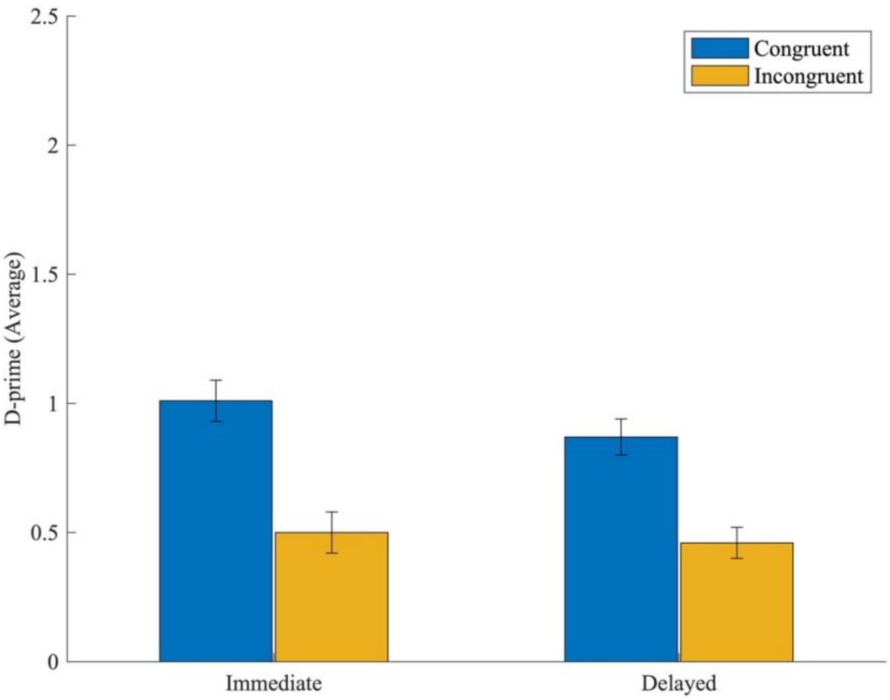

Figure 5.7 Average d’ scores for target memory in the Decontextualised condition at each test session (Immediate and Delayed) based on congruency (Congruent and Incongruent). Error bars represent ±1 SEM.

The results of the criterion analysis for the Decontextualised targets demonstrated that there were both main effects of *Congruency* ($F(1, 43) = 57.46$, $p < .001$, $\eta_p^2 = .57$) and *Time* ($F(1, 43) = 9.52$, $p = .004$, $\eta_p^2 = .18$). When the targets were presented in continually changing scenes, participants adopted a more conservative criterion threshold for Congruent targets ($M = .51$, $SE = .07$) than they did for Incongruent targets ($M = .11$, $SE = .06$). Participants additionally became more conservative as the length of time between study and test increased (Immediate: $M = .25$, $SE = .06$, Delayed: $M = .37$, $SE = .06$). Again, the *Congruency X Time* interaction ($F(1, 43) = .07$, $p = .790$, $\eta_p^2 < .01$) was found to not be significant (see Figure 5.8), and there was no significant difference between the *Nap* and *Wake* groups ($F(1, 43) = 1.77$, $p = .191$, $\eta_p^2 = .04$).

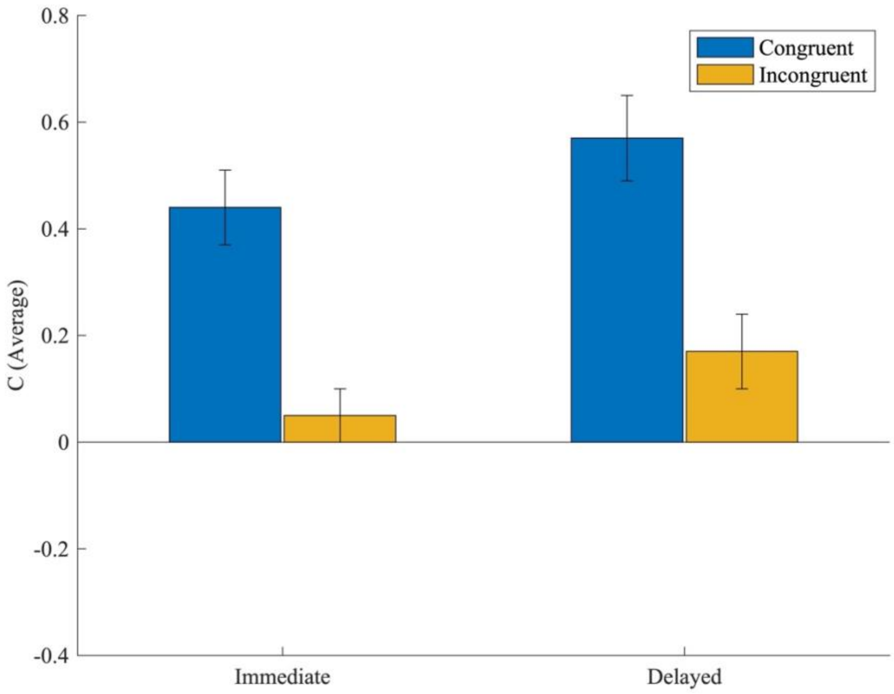

Figure 5.8 Average criterion for old/new decision in the Decontextualised condition at each test session (Immediate and Delayed) based on congruency (Congruent and Incongruent). Error bars represent ±1 SEM.

5.3.2 Sleep Questionnaires and Delay Conditions

Descriptive statistics from the sleep questionnaires that all participants completed, regardless of group, can be found in Table 5.2. Participant responses on the PSQI, the Epworth Sleepiness Scale and the Stanford Sleepiness Scale demonstrate that performance on the task was not influenced by how tired the participants were. An independent-samples t-test was conducted on the Stanford Sleepiness Scale score at the delayed test to ensure participants in the Wake group did not feel significantly more tired than those who had taken a nap. This revealed that subjective sleepiness did not significantly differ between the two delay groups ($t(43) = -.64$, $p = .523$).

Table 5.2 Average responses on the Pittsburgh Sleep Quality Index (PSQI, ranging from 1-21), Epworth Sleepiness Scale (ESS; ranging from 0-24) and the Stanford Sleepiness Scale (SSS; ranging from 1-7). Higher scores on the PSQI indicate poorer sleep quality. Higher scores on the ESS indicate excess sleepiness whilst lower scores on the SSS indicate less feelings of tiredness at that moment.

	N	Score (average)	SD
PSQI	45	4.24	± 1.37
Epworth Sleepiness Scale	45	5.58	± 3.33
Stanford Sleepiness Scale	45	3.19	± .08
SSS Encoding	45	2.38	± .09
SSS Immediate	45	4.38	± .97
SSS Delayed	45	3.19	± .08

Furthermore, there were no significant differences in d' scores between the sleep and wake groups in any of the three context conditions (all $p > .05$). Figure 5.9 shows performance at the Immediate and Delayed tests for the Sleep and Wake participants, collapsed across context and congruency conditions. This further demonstrates that there was no difference between the two groups.

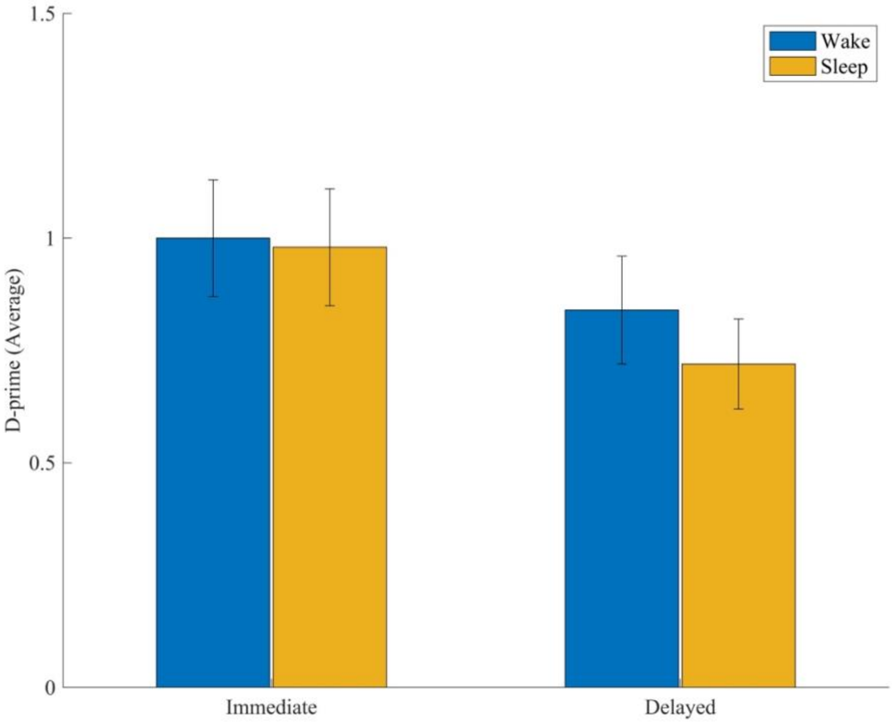

Figure 5.9 Average d' scores for target memory for the Wake and Sleep groups at each test session (Immediate and Delayed) collapsed across context and congruency conditions. Error bars represent ±1 SEM.

5.3.3 Nap Descriptives

Descriptive statistics for the naps can be found in Table 5.3.

Table 5.3 Descriptive statistics for the scored PSG data in the nap condition. Participants were given a 90-min sleep opportunity. Time in each stage of sleep is reported as average number of minutes and average percentage of the sleep period.

	N	Average (Min)	Average (%)
Total sleep time	24	79.81 (± 9.07)	
Wake after sleep onset	24	4.40 (± 8.68)	
Sleep efficiency	24		94.78 (± 9.61)
N1	24	15.94 (± 9.61)	21.01 (± 13.13)
N2	24	25.78 (± 13.00)	32.29 (± 14.90)
N3	24	23.00 (± 12.66)	29.54 (± 15.75)
REM	24	14.04 (± 11.32)	17.15 (± 13.68)

5.3.4 Behavioural Results Discussion

Analyses of the behavioural data fully replicated the results presented in Chapter 2 despite the experiment taking place in one day rather than across the span of a week. Semantically congruent targets repeatedly reinstated in the same scene-context (i.e., No Change condition) were found to be tightly bound to the scene-context in which they appeared which facilitated accurate retrieval, replicating the results from Chapter 2. This provides further evidence that object-scene binding occurs automatically (Hayes et al., 2007) and that strong associative links are formed between an item and the contextual details of an event. Tight object-scene binding still seems to be most evident for the congruent cases such that reinstating the studied scene-context at test led to significantly better memory for targets that were semantically congruent with the scene context. Conversely, it appears that the tight object-scene binding which resulted from presenting the target in the same scene each time it was encountered at encoding became a hinderance to memory performance at test when the scene-context was changed (i.e., Contextualised condition). The congruency advantage in this condition was abolished, with memory for both incongruent and congruent targets being severely impacted, providing further evidence for the idea that, at retrieval, the scene context no longer serves as a useful cue for either type of target.

Finally, encoding a target object in two different contexts breaks the object-scene binding, leading to a gist-like representation of the target (i.e., Decontextualised condition). Within this condition, a more abstract concept (schema) is being stored rather than the specific item being encoded. This allows the congruency advantage to emerge during retrieval due to the new scene context not proving useful for the incongruent targets at test. It is the fact that the congruent targets are associated with previously established schemas that allows the new contexts to provide a useful retrieval cue, leading to better performance for congruent targets in the Decontextualised condition. The present results provide further support for this argument, again showing significantly better performance for the congruent targets compared to incongruent targets. The behavioural results collectively demonstrate the robustness of the experimental paradigm and conclusions drawn in Chapter 2, yet the EEG results presented below provide insight into what neuronal processes underlie these effects.

5.3.5 EEG Results: Encoding

The primary frequency band of interest was the theta band (4-8 Hz). Neural activity for each trial was calculated, relative to pre-trial baseline activity, and grand averages for each condition were calculated. Dependent-samples t-tests were conducted using cluster-based permutation tests to determine whether there were significant differences in neural activity. As a reminder, Figure 5.10 shows the EEG channels and layout used in the present analyses.

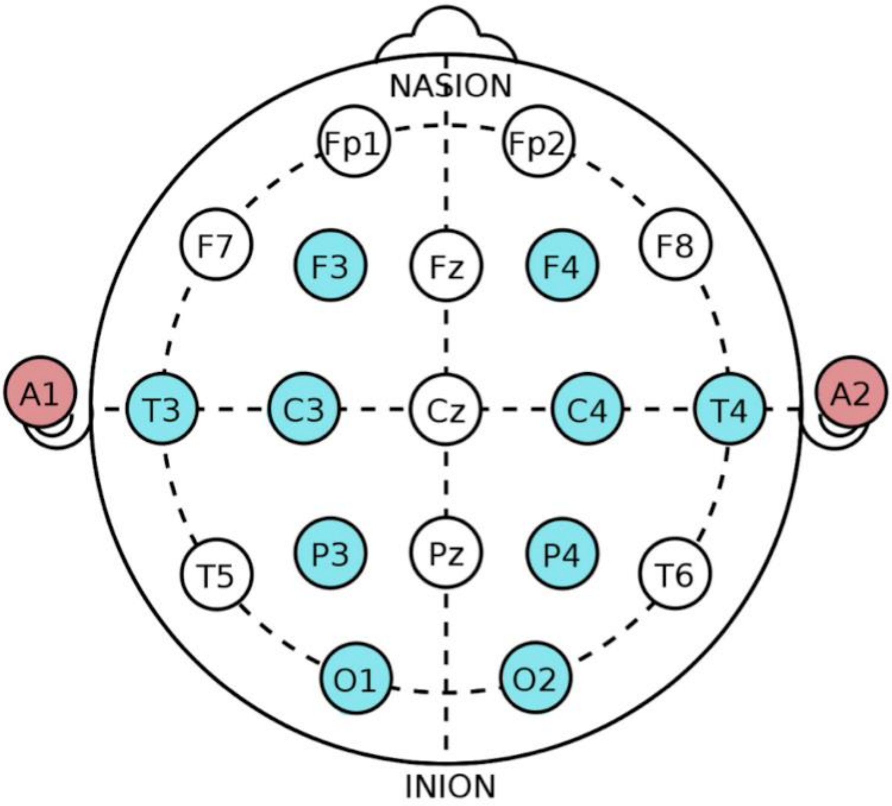

Figure 5.10 Diagram of EEG channels used in the study/analyses. Blue shaded channels represent channels included whereas the two red channels (i.e., A1 and A2) refer to the linked mastoids used as reference channels.

5.3.5.1 Presentation Order, Context and Congruency

The first step to understanding what neural patterns support encoding was to check for differences in activity during encoding. Three comparisons were made, respectively, between the first and second presentation of a target (i.e., Presentation 1 compared to Presentation 2), between targets in cases where the same encoding scenes are used and where different scenes are used across Presentation 1 and Presentation 2, and for Congruent compared to Incongruent targets. Note, the Same Encoding cases include the No Change and Contextualised conditions, the Different Encoding case is simply the

Decontextualised condition. The results demonstrated that there were no significant difference in theta activity between the first and second presentation (*p* > .05; see Figure 5.11a) or between the Same Encoding and Different Encoding cases (*p* > .05; see Figure 5.11b). However, there was a significant difference in theta activity based on congruency (see Figure 5.11c). Theta power significantly increased when a target was incongruent with the scene context relative to when the target was semantically congruent with the context (corresponding to two clusters). There was significantly more theta activity for all channels in the left hemisphere beginning ~.5-.95s after stimulus onset (*p* = .016) and for all channels in the right hemisphere beginning ~.5-.75s (*p* = .034).

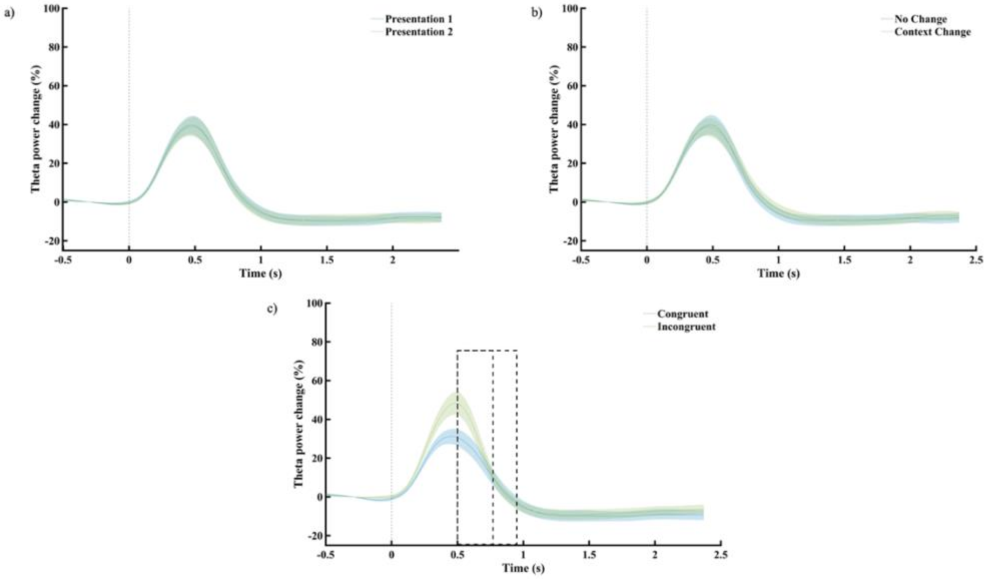

Figure 5.11 a) Average change in theta activity, relative to baseline, for the first presentation of a target versus the second presentation of a target. The shaded area around the lines represents ± 1 SEM. b) Average change in theta activity, relative to baseline, for targets presented in the same scene context versus targets presented in two different contexts. The shaded area around the lines represents ± 1 SEM. c) Average change in theta activity, relative to baseline, for congruent versus incongruent targets. Dotted boxes indicate the two significant clusters both beginning at ~.5s and lasting until ~.75s and ~.95s after stimulus onset. The shaded area around the lines represents ± 1 SEM.

5.3.5.2 Congruency and Context

Now critical comparisons were made for Congruent and Incongruent targets separately for the Same Encoding and Different Encoding cases. In both cases, data were averaged across Presentation 1 and Presentation 2, meaning any significant differences are carried by differences at the second presentation of the stimulus. For the Same Encoding cases there were no significant differences between congruent and incongruent targets in terms of theta activity (*p* > .05; see Figure 5.12a). However, for the Different Encoding cases there was a significant increase in theta activity for incongruent targets, compared to congruent, (corresponding to one significant negative cluster; see Figure 5.12b). In this condition, theta activity significantly increased in channels C4, P4 and T4 in the right hemisphere beginning ~.92-1.66s after stimulus onset (*p* = .014).

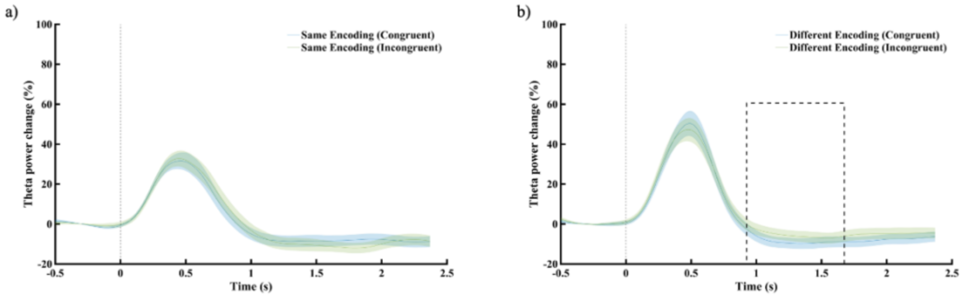

Figure 5.12 a) Average change in theta activity, relative to baseline, for congruent and incongruent targets in the Same Encoding condition (No Change and Contextualised targets). The shaded area around the lines represents ± 1 SEM. b) Average change in theta activity, relative to baseline, for congruent versus incongruent targets in the Different Encoding condition (Decontextualised targets). The dotted box indicates a significant cluster beginning ~.92-1.66s after stimulus onset. The shaded area around the lines represents ± 1 SEM.

5.3.6 EEG Results: Retrieval

5.3.6.1 Sleep does not Influence Neural Activity at Retrieval

The behavioural results demonstrated that there was no difference in memory performance between the Sleep and Wake groups. However, to ensure the delay conditions did not affect neural activity, theta activity at the immediate test (Test 1) was compared to the delayed test (Test 2) and between the Sleep and Wake groups. Note that Sleep/Wake was a between-subjects factor within the current paradigm, therefore there were 20 participants in each condition respectively. There were no significant differences in theta activity between Test 1 and Test 2 (no clusters identified; see Figure 5.13a) and there was no significant difference between the Sleep and Wake group at

Test 2 (*p* > .05; see Figure 5.13b). Therefore, the remaining analyses were collapsed across tests and groups.

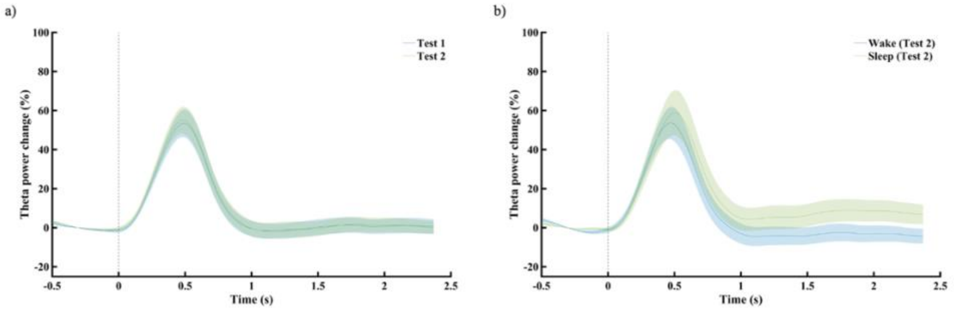

Figure 5.13 a) Average change in theta activity, relative to baseline, for Test 1 and Test 2. The shaded area around the lines represents ± 1 SEM. b) Average change in theta activity, relative to baseline, at Test 2 for participants in the Wake condition and participants in the Sleep condition. The shaded area around the lines represents ± 1 SEM.

5.3.6.2 Context and Congruency

Much like the findings from encoding, there was a significant difference in theta activity based on congruency (see Figure 5.14). Theta power significantly increased when a target was incongruent with the scene context relative to when the target was semantically congruent with the context (corresponding to two clusters). There was significantly more theta activity for all channels in the left hemisphere beginning ~.5-.79s after stimulus onset (*p* = .040) and for all channels in the right hemisphere beginning ~.5-.74s (*p* = .040). These results align with the changes in theta activity observed during encoding within the congruency condition (i.e., two significant clusters at similar time points during stimulus presentation).

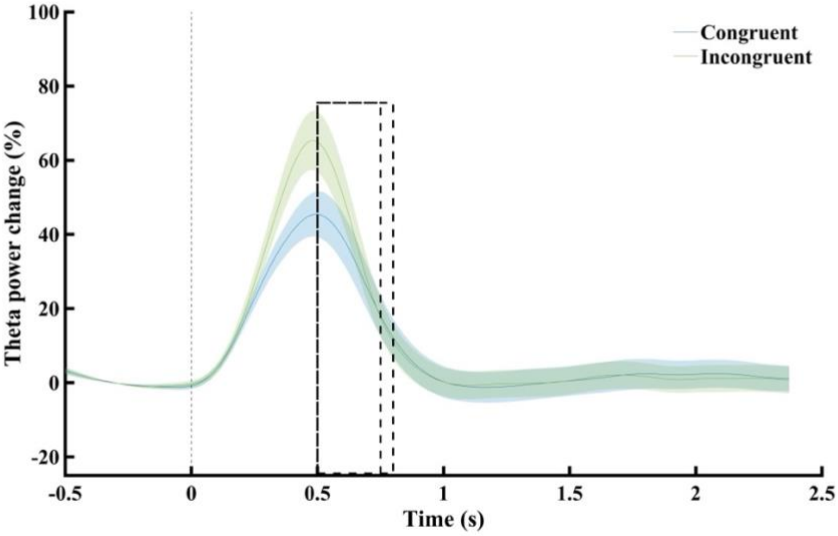

Figure 5.14 Average change in theta activity, relative to baseline, for congruent versus incongruent targets at test. The dotted boxes indicate two significant clusters, both beginning ~.5s after stimulus onset and lasting until ~.74 and .79s after stimulus onset. The shaded area around the lines represents ± 1 SEM.

Since theta activity was found to be sensitive to congruency during retrieval and to keep things consistent with the behavioural analyses, differences in theta activity for congruent and incongruent targets were calculated separately for each of the three context conditions (see Figure 5.15). There were no significant differences in theta activity between congruent and incongruent targets when the scene context was reinstated at test (i.e., No Change condition) or when there was a change in scene-context between study and test (i.e., Contextualised condition; all *p* > .05). However, when the target appeared in a third, unique scene context at test (i.e., Decontextualised condition), theta activity was significantly higher for incongruent targets compared to congruent targets (*p* = .022). This difference corresponded to one significant negative cluster, localised to all channels across the right hemisphere. Differences in theta were observed early in the trial presentation, ~.5-.88s after stimulus onset.

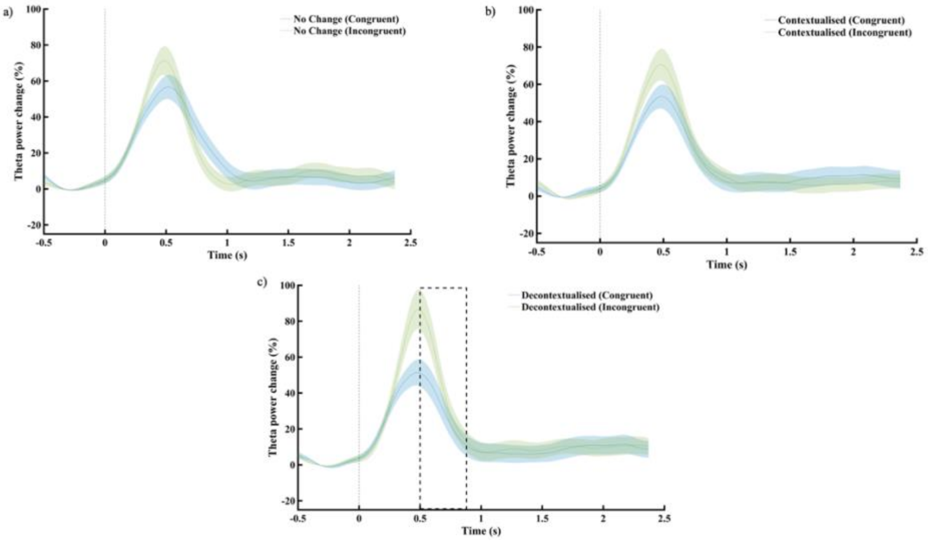

Figure 5.15 a) Average change in theta activity, relative to baseline, for congruent and incongruent targets appearing in the same context they were shown in during encoding (i.e., No Change). The shaded area around the lines represents ± 1 SEM. b) Average change in theta activity, relative to baseline, for congruent and incongruent targets presented in a new context at test after being presented in a repeated context during encoding (i.e., Contextualised). The shaded area around the lines represents ± 1 SEM. c) Average change in theta activity, relative to baseline, for congruent and incongruent targets presented in a third, new context at test after being presented in two different contexts during encoding (i.e., Decontextualised). The dotted box indicates a significant cluster beginning ~.5-.88s after stimulus onset. The shaded area around the lines represents ± 1 SEM.

5.3.7 EEG Results Discussion

The findings from the EEG data provided novel insight into what processes were implicated during encoding and retrieval. The key finding that emerged during both encoding and retrieval was that theta activity increased in response to incongruent information. This effect occurred early in the trial (~.5s), suggesting that theta activity is increased in response to perceptual processes, making participants quickly aware of the incongruent object within the scene and the semantic violations that were present. Furthermore, when looking at congruency effects within each of the context conditions it was only within the Decontextualised condition, again during both encoding and retrieval, that theta activity increased for the incongruent targets.

Perhaps the more interesting finding was that the effect of congruency within the Decontextualised condition differed between encoding and retrieval, despite both sessions demonstrating an increase in theta for the incongruent cases. The key difference within this condition was that, during encoding, the effect occurred late in the trial (~.92s) and during retrieval it occurred early in the trial (~.5s). This late time course of significant neuronal activity changes during encoding might be reflective of increased difficulty in the formation of a cohesive memory trace when the scene context changed on the second presentation of the target. A lack of episodic binding, in conjunction with a new scene context and an incongruent stimulus, makes forming a cohesive memory trace difficult. This contrasts with the changes in theta activity in response to an incongruent target in a third, new context at retrieval, with theta activity increasing as soon as perception of the event has registered (~.5s). It therefore appears to be the case that both encoding and retrieval of incongruent targets within the Decontextualisation was disrupted, a point which will be discussed in more detail in the General Discussion.

5.4 General Discussion

The present study aimed to investigate what neural mechanisms support the formation, consolidation and retrieval of visual episodic memories. Specifically, the role of theta during encoding and retrieval was examined. The study used a similar design to the previous three chapters and used the same stimuli as Chapter 2. Participants studied a set of target objects that were either semantically congruent or semantically incongruent with the scene context. Targets could appear in one of three context conditions in which the context remained consistent or changed. Rather than testing occurring across a week, participant spent an afternoon in the sleep lab where they completed the study phase and immediate test, spent a delay period either engaged in quite wakefulness or asleep, and then completed a delayed test. Neural recordings were obtained during encoding and tests. The behavioural results replicated the findings in Chapter 2. Better memory performance was observed for congruent targets in the No Change and Decontextualised conditions and there was no difference in performance between the congruent and incongruent targets in the Contextualised condition. Interestingly, there were no differences in performance or theta activity at the delayed test between the Sleep and Wake groups. However, theta activity was significantly increased during both encoding and retrieval when the target was incongruent with the scene context.

5.4.1 Context Reinstatement Aids Object Memory

Whilst the behavioural data suggests that the tight binding at encoding facilitated retrieval, replicating the results from Chapter 2, evidence from the neural data must be considered. Previous researchers have suggested that repeated encounters with an event reactivates the original memory trace, strengthening the memory trace via reconsolidation of the episode (Zöllner et al., 2023). Once reminded of a previous event, the previously formed memory trace becomes labile, leaving it susceptible to updating, changing or strengthening (Hupbach et al., 2007; Nadel et al., 2012; Winocur et al., 2009) It is possible that reactivation and reconsolidation occurred in the present study when the target was re-encountered, however there were no significant differences in theta activity between the first and second presentation during encoding. Increased theta activity has been associated with episodic memory formation (Atienza et al., 2011; Friese et al., 2013; Griffiths et al., 2019; Hanslmayr et al., 2012; Roux et al., 2022), item-context binding (Hanslmayr et al., 2009, 2016; Staudigl & Hanslmayr, 2013), and with associative binding (Köster et al., 2018). For example, Staudigl and Hanslmayr (2013) observed that theta power increases during encoding predicted successful retrieval when the study and text contexts matched and that theta power increases during encoding hindered memory when there was a change in context between study and test. Whilst the encoding data in the present study were not used to make predictions about retrieval and our targets were shown more than once during encoding, one might expect theta activity to increase within the No Change condition given the repetition of the scene-context during encoding if theta increases during encoding predict successful retrieval when there is a match in context. Within this condition there were no differences in theta activity between the congruent and incongruent targets during encoding, despite congruent targets being more accurately recognised. It seems the case that theta activity changes in response to context-reinstatement might be more specific to retrieval processes or subsequent memory effects rather than context-reinstatement alone. Previous research has further shown that theta oscillations, particularly over frontal electrodes, provide evidence of successful binding of an item to its context (Griffiths et al., 2021; Hanslmayr et al., 2009, 2016; Köster et al., 2018), yet the present results within the theta frequency do not elucidate whether the congruent targets were more tightly bound to their context during encoding.

One might wonder if the better performance observed for congruent targets was driven by neuronal processes at retrieval. Again, it appears that theta activity was not the underlying frequency band underpinning performance at retrieval for congruent targets in the No Change condition. Studies have shown that theta activity increases during retrieval for successfully remembered items (Hanslmayr et al., 2009, 2012; Heib et al., 2015; Osipova et al., 2006; Staudigl & Hanslmayr, 2013; Zhao et al., 2022). The present results demonstrated that there were no significant differences between congruent and incongruent targets, in terms of theta activity, during retrieval when the original context was reinstated. Despite not finding significant differences within the theta band for the No Change condition, it still appears to be the case that tight object-scene binding facilitates retrieval of the congruent targets when the scene-context is reinstated. It is possible that another frequency band may be contributing to this finding (see the Appendix for exploratory analyses), however which bands might be contributing was beyond the scope of the present study.

5.4.2 Changes in Context Hinders Memory

Like before, it was postulated that processes at encoding were driving the behavioural effects in the Contextualised condition - the tight object-scene binding resulted from encoding processes. The Contextualised targets were included in the Same Context condition for the encoding analyses due to the targets being presented twice in the same scene context. It was only during retrieval the context change occurred for this condition. Given that there were no significant differences in theta activity between congruent and incongruent targets during encoding when the target appeared in the same scene twice, it again seems that the behavioural results must be partially driven by processes at retrieval. Within the Contextualised condition, theta activity at retrieval did not differ significantly between the congruent and incongruent targets. Interference within the memory system is present at retrieval due to the target appearing in two consistent scenes during encoding, disrupting accurate recognition for the targets when they are presented in a new context at test. It remains unclear what frequency band might be contributing to the results (see Appendix for exploratory analyses and possible frequency band contribution), though it is clear that theta activity is not supporting this effect. Theta activity therefore, within the current paradigm, may be more sensitive to incongruent information.

5.4.3 Decontextualisation Leads to Less Dependence on Context

The behavioural results within the Decontextualised condition again replicated the findings from Chapter 2. Performance was significantly better for congruent targets within this condition, supporting the idea that a more abstract, gist-like concept of the target was being stored in memory rather than the specific target. It was assumed that processes at retrieval must be carrying these effects when considering the behavioural data alone, due to no object-scene binding occurring whilst encoding (i.e., the object never appeared in the same scene-context more than once). However, the EEG results demonstrated differences in theta activity during both encoding and retrieval.

There was a significant increase in theta activity for the incongruent targets compared to the congruent targets when the scene-context changed during encoding, carried by changes in activity at the second presentation, suggesting that neuronal processes during encoding may play some role in performance within this condition. Theta activity has been shown to increase in response to semantically incongruent stimuli (Dini et al., 2022; Schneider & Maguire, 2018) and in response to unexpected stimuli (Klimesch et al., 1994). The significant increase in theta activity observed during encoding when the scene-context changed, and the fact that target was incongruent is likely reflective of the unexpected nature of the target appearing in the new scene at the second presentation combined with the target violating semantic expectations of the scene.

Processes at retrieval must be contributing to performance in the Decontextualised condition given that associative binding was not possible during encoding due to the continually changing contexts. Indeed there was evidence that neuronal processes at retrieval contributed to performance as well. There was a significant increase in theta for the incongruent targets compared to the congruent targets in the Decontextualised condition at retrieval, early in the trial. This is likely reflecting the unexpected nature of the third new scene context in combination with the target continuing to violate semantic expectations about what should be present in the scene. Theta increases in this case might reflect increased processing needed to determine whether the target had been previously encountered, particularly with a new scene context that does not provide any helpful retrieval cues.

5.4.4 Theta Activity is Sensitive to Congruency and Changes in Context

It is of some note to find that congruency was the only factor that led to significant changes in theta activity during encoding and during retrieval. In both cases, significant increases in theta activity for incongruent targets was observed early in the trial (~.5-.95 Encoding; ~.5-.88s Retrieval) relative to congruent targets. This demonstrates that theta activity reflects the registration of incongruity very rapidly, presumably because of perceptual processes upon viewing the stimulus. Even more intriguing was the finding that, during both encoding and retrieval, there was a significant increase in theta activity in response to the incongruent targets only within the Decontextualised condition. During encoding, this effect is observed relatively late in the stimulus presentation (~.92-1.66s) whereas it occurs almost immediately during stimulus presentation at retrieval (~.5-.88s). This evidence can be viewed alongside the extant findings regarding theta activity implicated in verbal tasks. For instance, findings demonstrate an increase in theta in response to incongruent words (Schneider & Maguire, 2018), unrelated words (Maguire & Abel, 2013) and anomalies in the stimuli (Klimesch et al., 1994). However, it is important to note that the studies that have observed increases in theta activity in response to incongruent/anomalous stimuli have used different experimental paradigms in comparison to the present study, namely verbal stimuli as opposed to pictures. For example, Schneider and Maguire (2018) presented participants with 160 sentences that were either semantically correct or that contained a semantic error. Participants were asked to indicate whether the agent-action pairing was plausible and whether it was correct. Their findings demonstrated that there was increased theta activity in response to semantic errors (i.e., incongruent cases), localised to the left-frontal electrode sites. Furthermore, theta activity increases in the presence of anomalies due to increased encoding and retrieval demands (Klimesch et al., 1994), which is most likely what is contributing to the effects observed in the present study. Within the No Change and Contextualised condition, object-scene binding occurs due to the repetition of a consistent scene-context during encoding, though the strength of that binding varies based on the relationship of the target to the scene (i.e., congruent or incongruent). However, for the Decontextualised targets, there is a new scene-context presented each time the target is encountered. The new scene context combined with an incongruent or ‘anomalous’ target leads to difficulties in forming a coherent memory trace during encoding, reflected in the increased theta activity late in the trial for incongruent targets within this condition. This is further evidence by the increase in theta activity within this condition when participants were attempting to retrieve the target.

Recently, Alejandro et al. (2021) observed increases in theta activity when participants were retrieving incongruent words from word lists compared to when they were retrieving congruent words from the list. Our findings further support an increase in theta activity when retrieving incongruent visual stimuli. Theta activity during retrieval was only sensitive to congruency, with findings again demonstrating an increase in theta activity in response to incongruent stimuli during retrieval. As was the case during encoding, theta activity during retrieval was only significantly impacted by incongruent targets within the Decontextualised condition. This suggests that in addition to severely disrupting the formation of memories when a target was incongruent with the scene-context, the retrieval of such targets is also disrupted. Again, it could well be the case that this is partially driven by the increased demands of attempting to retrieve the target that has no pre-existing semantic associations with the scenes and the target being in a third, new context during retrieval.

One surprising finding was that there was no difference in theta activity when the context was reinstated. The current paradigm was, to the best of our knowledge, one of the first to examine neuronal responses to context reinstatement and context changes during visual episodic encoding. The literature demonstrating theta activity increases when the context at test matches the context at study (e.g., Staudigl & Hanslmayr, 2013) have primarily focused on verbal stimuli and used experimental designs that include a single presentation of a target during encoding prior to test. Perhaps the previous effects pertaining to theta and context reinstatement are specific to retrieval. Alternatively, Köster et al. (2018) put forth the argument that theta activity increases reflect successful binding of visual perceptual information (in their case binding between a pictured object and its background colour). Their study forced participants to form associations between a target and the background context, leading to intentional associations between the elements in each episode. It may well be the case that theta activity supports successful binding when the experimental task requires intentional encoding of the entire event. In the present experiment, participants were told to ignore the background context and to only remember the target object. Binding of a target with the background scene occurred automatically in the present study, and no evidence for theta's role in the process was observed. It was only when the context changed (i.e., Decontextualised condition) that changes in theta activity were observed such that, during encoding, increases in theta in response to incongruent targets were carried by power changes at the second presentation. Furthermore, when a third, new context was presented during retrieval, incongruent targets elicited a significant increase in theta activity. Within this condition no episodic binding was able to occur due to the scene-context changing each time the target was encountered, raising the question as to why increases in theta were observed in this case. This gives rise to the argument that theta activity is more sensitive to incongruent and/or anomalous stimuli given that this condition contained a surprising new context and an incongruent target each time the target was encountered. The incongruent and unexpected nature of the stimulus in conjunction with the new background scene leads to increased difficulty in forming and retrieving such memory traces; an effect theta is particularly sensitive to.

5.4.5 Sleep did not Influence Memory for Targets

It was surprising to find that there were no significant differences between the sleep and wake group, despite a large body demonstrating a memory benefit following a period of sleep. This was reflected in no significant differences in d' or theta activity between the sleep and wake groups. Previous studies have shown that sleep plays a beneficial role in memory when it occurs shortly after learning (Born et al., 2006; Gais et al., 2006). The beneficial effects of sleep on memory arise due to sleep facilitating consolidation of the newly encoded memory traces. These benefits have been observed across long periods of nocturnal sleep as well as in daytime naps of various lengths (see Lahl et al., 2008; Mander et al., 2011; Sweegers & Talamini, 2014; Tucker & Fishbein, 2008). Crucially, sleep influences the consolidation of memories across all modalities, but favours explicitly encoded material (Diekelmann & Born, 2010).

Two possible explanations for the lack of differences between the sleep and wake groups arise. Firstly, if consolidation during sleep favours explicitly encoded material (Diekelmann & Born, 2010), then it is possible that the implicit encoding of the background scene diminished the beneficial effects of sleep. Participants in the present study were asked to explicitly encode the target object. However, the behavioural results demonstrated that the background scene was automatically encoded alongside the target object despite instructions to ignore it. Alternatively, it is possible that the second presentation of the target established a strong enough memory trace that consolidation was not necessary for successful retrieval to occur. Further research is needed in order to establish what might be contributing to the lack of sleep effects. For example, one might consider manipulating the number of presentations of the target during encoding or manipulating whether the target appears as part of a scene or in isolation to determine whether the multiple presentations or explicit encoding diminished the beneficial effects of sleep.

5.5 Conclusion

The present study expanded the previous studies to examine how semantic congruency and context impact recognition for elements of visual episodic events. There was a particular focus on what neural mechanisms supported the congruency effects observed in object-scene binding and accurate visual recognition memory. The behavioural results further demonstrated that the strength of binding of elements in a visual episodic event contributes to memory for individual elements. When the same object-scene image was reinstated during encoding and at test then congruent targets were more accurately remembered than incongruent targets, with memory facilitated by the tight object-scene binding. When the scene context is repeated during encoding but changes between study and test, memory suffers, and the congruency memory advantage is abolished. Furthermore, when a congruent target appeared in different scene contexts of the same type at both study and test, the memory trace became more gist-like and schematic in nature, allowing a congruency advantage to re-emerge. Like in the previous studies, the scene at test matches with the stored gist representation and provides a benefit to memory. The same was not the case for incongruent targets because no gist “cues” are generated during encoding and therefore these do not operate at test. This was supported by an increase in theta activity for the incongruent cases during both encoding and retrieval. Furthermore, it was only within the Decontextualised condition that theta activity increased in response to incongruent targets, again during both encoding and retrieval. It appears to be the case that theta activity during encoding and retrieval is reflective of the neuronal processing of congruency information.

Appendix

Exploratory Analyses: Alpha Activity (8-12 Hz) During Retrieval

Exploratory analyses were conducted on the TFR data to examine whether alpha activity (8-12 Hz) supports accurate retrieval of the target objects within each of the three context conditions. Recent studies have shown that alpha activity is associated with memory performance (Alejandro et al., 2021; Packard et al., 2020) and that decreases within this frequency band are correlated with increases in performance (Griffiths et al., 2021). Furthermore, alpha power decreases have been linked to reactivation of specific features of a memory trace (Hanslmayr et al., 2012). Exploratory analyses within this frequency band were conducted in order to elucidate what neural mechanisms might underlie memory performance in the No Change and Contextualised condition.

Significant differences were observed within the alpha band, helping to shed light on what supports retrieval of the targets. There was a significant decrease in alpha activity, relative to baseline, for the congruent targets compared to the incongruent targets in the No Change condition (p = .004). The significant difference corresponded with one significant negative cluster, localised across the left hemisphere, and observed late in the trial presentation, ~1.09-2.06s after stimulus onset (See Figure 5.16a). The significant decrease in alpha activity for congruent cases in the No Change condition aligns with previous studies that demonstrate a decrease in alpha when performance increases and supports the congruency benefit observed in the present behavioural results. Interestingly, there were no significant differences in alpha activity between the congruent and incongruent targets in the Contextualised condition (p = .432), again supporting the behavioural findings in the current study (see Figure 5.16b). Finally, there was a significant decrease in alpha activity for the congruent targets compared to the incongruent targets within the Decontextualised condition (see Figure 5.16c). This corresponded to three significant negative clusters (p_1 = .010, p_2 = .024, p_3 = .030). Cluster 1 was located within the right hemisphere (F4, C4, P4, O2) and was observed late in the trial, ~1.46-2.35s after stimulus onset. Cluster 2 and Cluster 3 were both observed in the left hemisphere (F3, C3, T3) and were observed early in the trial, ~.5-1.06s after stimulus onset (Cluster 2), and late in the trial, ~1.96-2.35s after stimulus onset (Cluster 3).

Again, these findings align with the congruency benefit observed in the behavioural data in the Decontextualised condition.

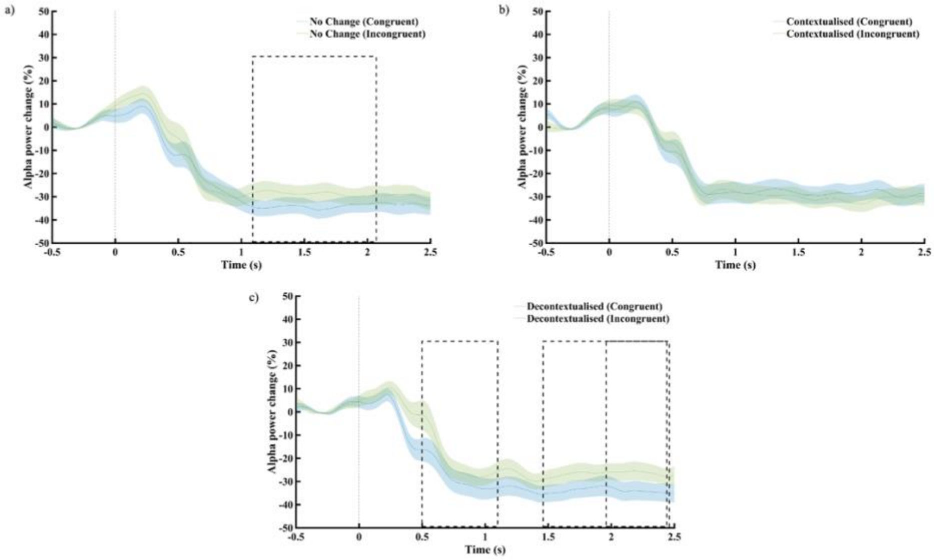

Figure 5.16 a) Average change in alpha activity, relative to baseline, for congruent and incongruent targets appearing in the same context they were shown in during encoding (i.e., No Change). The box indicates a significant cluster and the shaded area around the lines represents ± 1 SEM. b) Average change in alpha activity, relative to baseline, for congruent and incongruent targets presented in a new context at test after being presented in a repeated context during encoding (i.e., Contextualised). c) Average change in alpha activity, relative to baseline, for congruent and incongruent targets presented in a third, new context at test after being presented in two different contexts during encoding (i.e., Decontextualised). The boxes indicate significant clusters and the shaded area around the lines represents ± 1 SEM.

6. General Discussion

6.1 Overview of the Thesis

The aim of the present thesis was to better understand how context-binding and semantic congruency operate within visual episodic memory. Additionally, this thesis examined what, if any, processes are shared between visual and verbal episodic memory. Finally, this thesis sought to answer the question of whether effects of context and congruency emerge as a result of processes operating during encoding, retrieval or both. The experimental paradigm employed throughout the studies presented in this thesis utilised three main context conditions: targets presented in three consistent, reinstated contexts (i.e., No Change), targets presented in a reinstated context during encoding and in a new context at test (i.e., Contextualised) and targets appearing in three, unique changing scenes (i.e., Decontextualised). The relationship of the target to the context was manipulated in each of the experiments, with comparisons being made between targets that were semantically congruent or incongruent with the context (Chapters 2, 3 and 5) or targets that were animate or inanimate (Chapter 4).

The results in Chapter 2 demonstrated that, within visual episodic memory, the strength of binding of elements contributes to memory for the individual elements. This results in both a memory benefit and memory decrement depending on how the target is presented at test. For example, when the targets are semantically congruent with the scene-context and the context is reinstated at test, memory is aided. However, when the scene-context changes at test and the target is semantically congruent, memory for the target is hindered. Target objects were tightly bound to the scene-context they were presented in during encoding due to the targets being presented twice in the same scene. This allowed a congruent benefit to emerge when the context was further reinstated at test but became a hinderance to memory when the context changed at test (i.e., congruency advantage abolished in the Contextualised condition). It was concluded, due to the behavioural nature of the experiment, that processes at encoding and/or retrieval could be supporting accurate retrieval. It was only in the Decontextualised condition that one could confidently infer that processes at retrieval were contributing to performance. Within this condition, a congruency advantage re-emerged most likely due to a more schematic representation of the target being stored (i.e., a lamp in a living room rather than a specific lamp). This leads to congruent targets being better remembered than incongruent targets as the previously formed schemas for the scenes aid memory for a gist-like representation of the target. No such memory advantage appears for the incongruent targets at test due to the cues generating no helpful retrieval associations at test.

Researchers suggested that verbal and visual memory share processing characteristics (Standing, 1973) and may activate common networks during encoding and retrieval (Hamm et al., 2002). It has been well established that a congruency benefit is observed within the verbal domain (Bein et al., 2015; Höltje et al., 2019; Packard et al., 2017; Tibon et al., 2017; van Kesteren et al., 2020). This therefore led to the study presented in Chapter 3 whereby the congruency and context manipulations employed in the visual study presented in Chapter 2 were used to test verbal episodic memory. Here it was observed that the congruency differences were abolished in all three of the context conditions suggesting that participants were able to dissociate the targets from the contexts within verbal memory. This was contrary to the holistic and automatic encoding of the scene-contexts observed when using pictorial stimuli.

In addition to a consistent congruency advantage observed in verbal memory studies, an advantage of animate (i.e., living things) stimuli has been widely observed within the verbal domain (e.g., Bonin et al., 2014, 2015; Gelin et al., 2018; Hovhannisyan et al., 2021; Nairne et al., 2013, 2017). This, in conjunction with the results from Chapter 3, led to the examination of whether the animacy effects observed in the wider verbal literature would translate to the visual domain within the contextual manipulations of the previous thesis paradigms. An interesting pattern of results emerged in this case. It was the case that the pattern of effects observed in the data were driven by performance within the inanimate cases. This was evidenced by relatively stable recognition scores for the animate targets across the three context conditions, numerically speaking. This strengthens the argument that target objects with an expected, fixed location are tightly bound and associated with their scene context, making the holistic encoding difficult to overcome. Whilst there was significantly better performance for the animate targets when the object was studied in one context and tested in a new context, this effect was carried by particularly poor scores for the inanimate targets. The results, as a whole, point to the animate targets being less scene specific, making them less tightly bound to a scene during encoding and less susceptible to interference when there is a change in scene-context. This finding shed light on the paradigm-specific nature of animacy effects, supporting the argument put forth by Schreiner et al. (2013) that task requirements influence whether or not an animacy effect will emerge in experimental results. Nevertheless, the findings in Chapter 4 lend credence to the argument that elements in a visual episode are difficult to dissociate from the complete episode, leading to both a hinderance and supportive role of context to specific elements of the event at retrieval.

Finally, the study presented in Chapter 5 returned to the question of whether processes during encoding or retrieval supported the results observed in the previous chapters. The design and stimuli used mirrored those used in Chapter 2 but included EEG recordings throughout the experiment. Of particular interest was neuronal activity within the theta band (4-8 Hz; Iber et al., 2007) which has been previously shown to promote episodic memory formation (Atienza et al., 2011; Friese et al., 2013; Griffiths et al., 2019; Hanslmayr et al., 2012; Roux et al., 2022), item-context binding (Hanslmayr et al., 2009, 2016; Staudigl & Hanslmayr, 2013), and with associative binding (Köster et al., 2018). Changes in theta activity have also been observed in response to incongruent (e.g., Dini et al., 2022; Schneider & Maguire, 2018) and anomalous (Klimesch et al., 1994) information. The behavioural results replicated the findings in Chapter 2. It was only in response to congruency that differences in theta activity were observed, such that theta activity significantly increased in response to the incongruent targets at both encoding and test. Furthermore, it was only within the Decontextualised condition that there were differences in theta activity based on congruency during both encoding and retrieval.

The effect of congruency on theta activity during both encoding and retrieval occurred early in the trial, starting at roughly .5s at both sessions. Not only does this support previous literature demonstrating theta’s sensitivity to incongruent and anomalous information, but it appears to be the case that expectations of what should be present in a scene occurs early in the visual experience, activating the relevant schemas associated with the scene. This was particularly apparent within the Decontextualised condition, where theta activity was significantly increased for the incongruent targets again both during encoding and retrieval. Increases in theta activity within this condition occurred again at roughly .5s during retrieval, providing further evidence of the early detection of schematic violations. However, this increase in theta activity occurred relatively late in the trial during encoding, beginning at roughly .92s after stimulus onset, and was carried by theta activity changes during the second presentation of the stimulus. This late time course of significant neuronal activity changes begins to shed light on the particularly poor performance observed in this condition due to an increased difficulty in the formation of a cohesive memory trace. The argument made throughout the behavioural studies was that no object-scene binding was able to take place within the Decontextualised condition due to the scene changing each time the object was encountered. No episodic binding taking place in conjunction with a new scene context and an incongruent stimulus makes forming a cohesive memory trace difficult, evidenced by the poor performance behaviourally and the increase in theta activity late in the trial and carried by changes in neuronal activity on the second presentation of the target. Previous research has demonstrated that theta activity will increase when a participant is presented with unexpected stimuli (Klimesch et al., 1994), a factor that is likely contributing to the later time course of events within this condition during encoding combined with the increased difficulty of forming a memory trace for the incongruent cases. Whilst the study presented in Chapter 5 was unable to find evidence for a role of theta in episodic memory formation or context-binding to align with previous research in the area, it strengthened the evidence for theta's sensitivity to congruency and shed light on the potential neuronal mechanisms driving the behavioural effects observed throughout the thesis.

6.2 Implications of Findings

6.2.1 Expanding the findings of Evans and Wolfe (2022)

One of the aims of the thesis was to examine and expand the studies presented in a series of experiments by Evans and Wolfe (2022). They presented 12 experiments in which they examined the negative role of context and under what conditions such an effect arises. One of the key arguments they made was "that tight, automatic associations between target items and sensible backgrounds are crowding or hiding the mnemonic representation of those targets" (p. 238; Evans & Wolfe, 2022). At a very general level, a similar pattern of effects were observed across our experiments such that changing the scene-context between study and test led to particularly poor memory for target objects, though this effect is not likely driven by crowding effects. It appears to be the case that schemas and pre-existing semantic expectations about what should be present in a scene is a more accurate predictor of how context and congruency will support accurate retrieval.

The conclusions which can be drawn from the findings in Chapter 2 suggest that, at least in terms of congruency, the results from Evans and Wolfe (2022) did not replicate. They claim that memory was better for incongruent objects appearing in real-world scenes, yet the statistical findings from both Chapters 2 and 5 support a congruency advantage, particularly within the framework of reinstating a background scene and when the scene consistently changes, putting the current findings in alignment with the congruency advantage observed in the wider literature. The congruency advantage that emerges in the present paradigm is robust, evidenced by the findings replicating in Chapter 5 and providing further evidence against the incongruent benefit argued by Evans and Wolfe (2022). They do not dwell on the incongruent benefit observed and instead focus the final experiments on what happens when a target is presented multiple times during encoding and then presented at test in either the same background or in a new background. They utilised an experimental design similar to the design employed in the present thesis where participants were presented with target objects twice during encoding and memory for the target was tested once. Target objects were presented in a reinstated context during encoding and in a new context during test (i.e., Experiment 10B; our Contextualised condition) or were presented in a new scene each time it was encountered (i.e., Experiment 10A; our Decontextualised condition). However, no consideration for incongruent cases was taken in either Experiment 10A or 10B, they only used targets that were semantically congruent with the background. This left a gap in the literature, one that the current research was able to fill. It appears that a congruency advantage emerges under specific contextual manipulation conditions (e.g., context reinstatement and decontextualisation).

Perhaps the more interesting question that was not addressed in their study was whether the effects observed were carried by processes operating during encoding, retrieval or both. This became a main focus of the present thesis, with evidence being examined through behavioural and neuroscientific measures. The main pattern of findings that emerged from the behavioural data was that, within the decontextualised condition, the effects must be carried by processes operating during retrieval. This was due to the target objects having no opportunity to bind to a context during encoding (i.e., the context continually changed). It was therefore at retrieval when effects operated; either the new context associated schematically related information to support retrieval or continued to challenge existing schemas, hindering retrieval. The locus of effect within the no change and contextualised conditions were not clearly evidenced by the behavioural results in any of the studies, however. It was only in Chapter 5 that EEG evidence started to shed light on where these effects operate. Evidence for processes operating at both encoding and retrieval were found within the decontextualised condition (discussed in detail below).

One final way the current thesis expanded the findings of Evans and Wolfe (2022) was by examining whether visual episodic memory would be sensitive to animacy effects. Evans and Wolfe (2022) argued that the tight object-scene binding in their studies was driven by the fact that the target objects used do not move of their own accord, enforcing a “standard representation” of the scene context. This led to the argument that if the objects could appear anywhere in a scene, then tight object-scene binding would not occur. The theory that tight object-scene binding would not occur for targets that could appear anywhere in a scene (e.g., cats, dogs, humans) was tested in Chapter 4. There is a fair amount of previous research that demonstrates superior memory for animate targets compared to inanimate targets (e.g., Bonin et al., 2014, 2015; Gelin et al., 2018; Hovhannisyan et al., 2021; Nairne et al., 2013, 2017; Schreiner et al., 2023; Sun et al., 2011; VanArsdall et al., 2013). However, these animacy effects have been established within the verbal domain and a very limited number of studies have examined the effect using images (see Sun et al., 2011 and Bonin et al., 2014; novel contributions of the current study will be discussed in the next section). The results presented in Chapter 4 supported the argument made by Evans and Wolfe (2022) that animate objects were not as tightly bound to their scene-contexts. The fact that performance remained stable for the animate targets across the three context conditions suggests that the movability and lack of fixed position for the animate targets leads to the targets being less tightly bound to the background scene. For example, you will consistently find a kettle on a kitchen counter whereas you might find a cat in the kitchen on the floor, table or counters. This, therefore, leads to less reliance on a specific scene-context at retrieval and makes them less susceptible to interference when the scene-context changes, contrary to the inanimate targets which are always found in a fixed location and are tightly bound to the scene in which they are encountered.

The present thesis has taken the 12 experiments conducted by Evans and Wolfe (2022) and expanded on their findings to better understand how congruency and context influences visual episodic memory. More importantly, novel paradigms have been used to test well-established theories, providing further insights into the field of memory research. The next section will address the novel contributions to the field in more detail.

6.2.2 Novel Contributions to the field

The first key novel contribution of the thesis can be observed in the study presented in Chapter 3. The verbal episodic memory study was the first to use sentences in a memory paradigm examining context binding. This is not to say that this was the first study to use sentences in an episodic memory paradigm nor in the examination of classic congruency effects, rather it was the first to examine how congruency operates within the framework of binding to a sentence context. Previous studies using sentence stimuli to examine congruency effects typically presented participants with sentences one word at a time (Besson et al., 1992; Haeuser & Kray, 2023; Kutas & Hillyard, 1980) and tested memory for targets either as a cued recall task (Besson et al., 1992) or by having participants make old/new judgements for presentations of only the target word at test (Haeuser & Kray, 2023). The closest study to employ methods similar to the present study was Flegal et al. (2014) in which entire sentences were presented to participants during encoding. However, each sentence contained a congruent and incongruent target word. The old/new task at test presented participants with a single word on the screen as a cue followed by a single word that they judged as old/new (Flegal et al., 2014) rather than placing the target word back into a sentence context. More importantly, the context in these studies is established by only presenting the target word once during encoding in one sentence context. In each of the aforementioned studies, results point to better memory for congruent targets. This aligns with the wider literature that supports a congruency advantage when faced with linguistic stimuli (see Chapter 3 for a full review of the literature). However, the present thesis demonstrated that congruency differences can be abolished under specific testing conditions within the verbal episodic memory domain.

The verbal study presented in Chapter 3 was not the only chapter to provide novel contributions to the field. The animacy study presented in Chapter 4 was also the first to examine animacy effects in visual long-term memory. Animacy effects were also well established within the verbal memory domain (e.g., Bonin et al., 2014, 2015; Gelin et al., 2018; Nairne et al., 2013; Schreiner et al., 2023; VanArsdall et al., 2013; see Chapter 4 for a full review of the literature). It was noted that a very limited number of studies have examined animacy effects within the visual domain, with an emphasis on visual attention and change detection tasks (Bonin et al., 2014; New et al., 2007; Sun et al., 2011). The study presented in Chapter 4 was, to the best of our knowledge, the first to examine animacy effects in visual episodic memory and in relation to context-binding. It was demonstrated that the animacy advantage can be abolished under certain encoding and retrieval conditions, dependent upon performance in the inanimate condition and the specific requirements of the experimental paradigm. These findings challenged the established argument that the human memory system selectively solves ‘critical adaptive problems’ related to survival needs, giving priority to animate stimuli. Perhaps this argument remains the case when solely examining visual attention, but to extend the animate monitoring hypothesis to visual long-term memory has its challenges based on the present findings.

Schreiner et al. (2023) was the only study to examine animacy effects in terms of an episodic binding paradigm, though their study examined the effect within the verbal domain. They argued that an animate target appearing in a sentence acts as a binding agent by providing a grammatical subject and structural awareness that creates a cohesive episodic event for the sentence. However, as was discovered across Chapters 2-4, verbal and visual episodic memory systems differ in the way they operate. Whilst an animate target might be needed to facilitate a grammatically complete event representation in verbal memory, this does not appear to be the case for pictorial stimuli.

The rich contextual details offered by pictures are difficult to overcome in memory, with the ability to separate an object from its background scene proving difficult (see Chapter 2). However, when the task requires remembering a specific word from a sentence, it is easier to filter out the unnecessary contextual details and to remember just the word (see Chapter 3). It is therefore argued that the controlled visual and verbal memory studies presented in Chapters 2-4 provide further novel contributions to the field. Recall that the studies presented in this thesis were all designed to implement the same experimental paradigm, allowing general theoretical comparisons across studies and types of stimuli. For example, the verbal contextual binding study presented in Chapter 3 was designed to be a verbal replication and extension of the visual study presented in Chapter 2. Indeed we see that the paradigm and results presented in this thesis are robust, with results perfectly replicating between Chapters 2 (behavioural) and 5 (EEG) when the same stimuli and design were used. We can therefore conclude that a main novel finding from the thesis, as a whole, is that episodic binding differs between visual and verbal episodic memory in terms of how people are able to dissociate a target from its context. This sheds light on the importance of clearly distinguishing between modalities when making claims about how memory operates.

Finally, this thesis was the first to examine how context repetitions/changes and congruency influences memory during encoding, particularly when considering the neural signatures of successful encoding (i.e., Chapter 5). Extensive previous research, focused primarily on verbal stimuli and single presentations of a target during encoding, demonstrated that theta activity (4-8 Hz) plays a key role in episodic memory formation (Atienza et al., 2011; Friese et al., 2013; Griffiths et al., 2019; Hanslmayr et al., 2012; Roux et al., 2022; see Chapter 5 for a full review of the literature). Moreover, theta activity increases when the study and test context matches, driving successful episodic recollection via retrieval of contextual details (Staudigl & Hanslmayr, 2013). In Chapter 5, the only significant changes in theta activity were observed in response to congruity (another well-established factor shown to increase theta activity). More specifically, it was only within the Decontextualised condition that differences in theta activity were observed, namely that there was an increase in theta activity for the incongruent targets during both encoding and retrieval. As this was the first study to examine changes in theta activity to multiple presentations of a stimulus during encoding, it was surprising to find that the repeated presentation in a reinstated context was not marked by changes in theta activity, consistent with previous study/test reinstatement effects. However, given that there was an increase in theta activity within the Decontextualised condition at encoding, the effects were carried by the second presentation of the stimulus late in the trial period, providing new evidence that incongruent stimuli are more difficult to integrate into cohesive memory traces which has detrimental consequences at retrieval.

6.3 Concluding Remarks

This thesis examined some of the factors that influence the formation, consolidation and retrieval of visual episodic memories with a particular emphasis on congruency and context-binding. Further, it examined what neuronal mechanisms underlie such processes. Classic theoretical effects were challenged (e.g., congruency, animacy, theta activity) within the scope of novel experimental paradigms to provide insight into the way in which episodic memories operate. This highlights the importance of more nuanced paradigms that are more naturalistic in order to garner a comprehensive investigation of the way the brain makes sense of the complex visual world around us. Considering the example of student exam performance in the introduction, perhaps the way students are able to overcome the detrimental effects of changes in context concerns the more verbal nature of the revised content. Results in the present thesis would suggest that the changes in learning/study/exam context are overcome by an ability to dissociate the target information from the unnecessary contextual details surrounding the event. However, if the material is more visual in nature, these effects of contexts are much more difficult to overcome due to the holistic encoding of the visual event. It can be concluded that context interacts with other factors present during the encoding and retrieval of an event, supported by differing neuronal markers to achieve the goal of accurate retrieval.

References

Alejandro, R. J., Packard, P. A., Steiger, T. K., Fuentemilla, L., & Bunzeck, N. (2021). Semantic Congruence Drives Long-Term Memory and Similarly Affects Neural Retrieval Dynamics in Young and Older Adults. *Frontiers in Aging Neuroscience*, *13*, 683908. https://doi.org/10.3389/fnagi.2021.683908

Andermane, N., & Bowers, J. S. (2015). Detailed and gist-like visual memories are forgotten at similar rates over the course of a week. *Psychonomic Bulletin & Review*, *22*(5), 1358–1363. https://doi.org/10.3758/s13423-015-0800-0

Anderson, M. C., & Hulbert, J. C. (2020). Active Forgetting: Adaptation of Memory by Prefrontal Control. *Annual Review of Psychology*, *72*(1). https://doi.org/10.1146/annurev-psych-072720-094140

Anderson, R. C. (1984). Role of the Reader's Schema in Comprehension, Learning, and Memory. In *Learning to Read in American Schools: Basal Readers and Content Texts* (pp. 243–257). Lawrence Erlbaum Associates.

Atienza, M., Crespo-Garcia, M., & Cantero, J. L. (2011). Semantic Congruence Enhances Memory of Episodic Associations: Role of Theta Oscillations. *Journal of Cognitive Neuroscience*, *23*(1), 75–90. https://doi.org/10.1162/jocn.2009.21358

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory.* (Vol. 2, pp. 89-195). New York: Academic Press.

Baddeley, A. (1999). *Essentials of Human Memory* (Classic Edition). Psychology Press.

Baddeley, A. D. (1982). Domains of Recollection. *Psychological Review*, *89*(6), 708–729.

Bein, O., Livneh, N., Reggev, N., Gilead, M., Goshen-Gottstein, Y., & Maril, A. (2015). Delineating the Effect of Semantic Congruency on Episodic Memory: The Role of Integration and Relatedness. *PLOS ONE*, *10*(2), e0115624. https://doi.org/10.1371/journal.pone.0115624

Berron, D., Schutze, H., Maass, A., Cardenas-Blanco, A., Kuijf, H. J., Kumaran, D., & Duzel, E. (2016). Strong Evidence for Pattern

Separation in Human Dentate Gyrus. *Journal of Neuroscience*, *36*(29), 7569–7579. https://doi.org/10.1523/JNEUROSCI.0518-16.2016

Besson, M., Kutas, M., & Petten, C. V. (1992). An Event-Related Potential (ERP) Analysis of Semantic Congruity and Repetition Effects in Sentences. *Journal of Cognitive Neuroscience*, *4*(2), 132–149. https://doi.org/10.1162/jocn.1992.4.2.132

Bloom, P. A., & Fischler, I. (1980). Completion norms for 329 sentence contexts. *Memory & Cognition*, *8*(6), 631–642. https://doi.org/10.3758/BF03213783

Bonin, P., Gelin, M., & Bugaiska, A. (2014). Animates are better remembered than inanimates: Further evidence from word and picture stimuli. *Memory & Cognition*, *42*(3), 370–382. https://doi.org/10.3758/s13421-013-0368-8

Bonin, P., Gelin, M., Laroche, B., Méot, A., & Bugaiska, A. (2015). The “How” of Animacy Effects in Episodic Memory. *Experimental Psychology*, *62*(6), 371–384. https://doi.org/10.1027/1618-3169/a000308

Born, J., Rasch, B. & Gais, S. (2006). Sleep to Remember. *The Neuroscientist*, *12*(5), 410-424. https://doi.org/10.1177/1073858406292647

Born, J., & Wilhelm, I. (2012). System consolidation of memory during sleep. *Psychological Research*, *76*(2), 192–203. https://doi.org/10.1007/s00426-011-0335-6

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences*, *105*(38), 14325–14329. https://doi.org/10.1073/pnas.0803390105

Brady, T. F., Konkle, T., Gill, J., Oliva, A., & Alvarez, G. A. (2013). Visual Long-Term Memory Has the Same Limit on Fidelity as Visual Working Memory. *Psychological Science*, *24*(6), 981–990. https://doi.org/10.1177/0956797612465439

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Brewer, W. F., & Treyens, J. C. (1981). Role of schemata in memory for places. *Cognitive Psychology*, *13*(2), 207–230. https://doi.org/10.1016/0010-0285(81)90008-6

Brooks, D. I., Rasmussen, I. P., & Hollingworth, A. (2010). The nesting of search contexts within natural scenes: Evidence from contextual cuing. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(6), 1406–1418. https://doi.org/10.1037/a0019257

Brunellière, A., Delrue, L., & Auran, C. (2020). The contribution of audiovisual speech to lexical-semantic processing in natural spoken sentences. *Language, Cognition and Neuroscience*, *35*(6), 694–711. https://doi.org/10.1080/23273798.2019.1641612

Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Research*, *28*(2), 193–213. https://doi.org/10.1016/0165-1781(89)90047-4

Cairney, S. A., Durrant, S. J., Musgrove, H., & Lewis, P. A. (2011). Sleep and environmental context: Interactive effects for memory. *Experimental Brain Research*, *214*(1), 83–92. https://doi.org/10.1007/s00221-011-2808-7

Castelhano, M., & Henderson, J. (2005). Incidental visual memory for objects in scenes. *Visual Cognition*, *12*(6), 1017–1040. https://doi.org/10.1080/13506280444000634

Cohen, M. X. (2014). *Analyzing Neural Time Series Data: Theory and Practice*. The MIT Press.

Cox, R., Tijdens, R. R., Meeter, M. M., Sweegers, C. C. G., & Talamini, L. M. (2014). Time, Not Sleep, Unbinds Contexts from Item Memory. *PLoS ONE*, *9*(2), e88307. https://doi.org/10.1371/journal.pone.0088307

Craig, M., Knowles, C., Hill, S., & Dewar, M. (2021). A study on episodic memory reconsolidation that tells us more about consolidation. *Learning & Memory*, *28*(2), 30–33. https://doi.org/10.1101/lm.052274.120

Criss, A. H., & Howard, M. W. (2015). *Models of Episodic Memory* (J. R. Busemeyer, Z. Wang, J. T. Townsend, & A. Eidels, Eds.; Vol. 1). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199957996.013.8

Davenport, J. L., & Potter, M. C. (2004). Scene Consistency in Object and Background Perception. *Psychological Science*, *15*(8), 559–564. https://doi.org/10.1111/j.0956-7976.2004.00719.x

Deason, R. G., Hussey, E. P., Ally, B. A., & Budson, A. E. (2012). Changes in response bias with different study-test delays: Evidence from young adults, older adults, and patients with Alzheimer's disease. *Neuropsychology*, *26*(1), 119–126. https://doi.org/10.1037/a0026330

DeLong, K. A., & Kutas, M. (2020). Comprehending surprising sentences: Sensitivity of post-N400 positivities to contextual congruity and semantic relatedness. *Language, Cognition and Neuroscience*, *35*(8), 1044–1063. https://doi.org/10.1080/23273798.2019.1708960

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, *8*(8), 1117–1121. https://doi.org/10.1038/nn1504

Denis, D., Bottary, R., Cunningham, T. J., Tcheukado, M.-C., & Payne, J. D. (2023). The influence of encoding strategy on associative memory consolidation across wake and sleep. *Learning & Memory*, *30*(9), 185–191. https://doi.org/10.1101/lm.053765.123

Denis, D., Mylonas, D., Poskanzer, C., Bursal, V., Payne, J. D., & Stickgold, R. (2021). Sleep Spindles Preferentially Consolidate Weakly Encoded Memories. *The Journal of Neuroscience*, *41*(18), 4088–4099. https://doi.org/10.1523/JNEUROSCI.0818-20.2021

Denis, D., Schapiro, A. C., Poskanzer, C., Bursal, V., Charon, L., Morgan, A., & Stickgold, R. (2020). The roles of item exposure and visualization success in the consolidation of memories across wake and sleep. *Learning & Memory*, *27*(11), 451–456. https://doi.org/10.1101/lm.051383.120

Diekelmann, S. & Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience*, *11*(2), 114-126. https://doi.org/10.1038/nrn2762

Dini, H., Simonetti, A., Bigne, E., & Bruni, L. E. (2022). EEG theta and N400 responses to congruent versus incongruent brand logos. *Scientific Reports*, *12*(1), 4490. https://doi.org/10.1038/s41598-022-08363-1

Dudai, Y. (2004). The Neurobiology of Consolidations, Or, How Stable is the Engram? *Annual Review of Psychology*, *55*(1), 51–86. https://doi.org/10.1146/annurev.psych.55.090902.142050

Eich, E. (1985). Context, Memory, and Integrated Item/Context Imagery. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*(4), 764–770. https://doi.org/10.1037/0278-7393.11.1-4.764

Evans, K. K., & Baddeley, A. (2018). Intention, attention and long-term memory for visual scenes: It all depends on the scenes. *Cognition*, *180*, 24–37. https://doi.org/10.1016/j.cognition.2018.06.022

Evans, K. K., & Wolfe, J. M. (2022). Sometimes it helps to be taken out of context: Memory for objects in scenes. *Visual Cognition*. https://doi.org/10.1080/13506285.2021.2023245

Flegal, K. E., Marín-Gutiérrez, A., Ragland, J. D., & Ranganath, C. (2014). Brain Mechanisms of Successful Recognition through Retrieval of Semantic Context. *Journal of Cognitive Neuroscience*, *26*(8), 1694–1704. https://doi.org/10.1162/jocn_a_00587

Fogel, S. M., Smith, C. T., & Beninger, R. J. (2009). Evidence for 2-stage models of sleep and memory: Learning-dependent changes in spindles and theta in rats. *Brain Research Bulletin*, *79*(6), 445–451. https://doi.org/10.1016/j.brainresbull.2009.03.002

Friedman, A. (1979). Framing Pictures: The Role of Knowledge in Automatized Encoding and Memory for Gist. *Journal of Experimental Psychology: General*, *108*(3), 316–355. https://doi.org/10.1037//0096-3445.108.3.316

Friese, U., Köster, M., Hassler, U., Martens, U., Trujillo-Barreto, N., & Gruber, T. (2013). Successful memory encoding is associated with increased cross-frequency coupling between frontal theta and posterior gamma oscillations in human scalp-recorded EEG. *NeuroImage*, *66*, 642–647. https://doi.org/10.1016/j.neuroimage.2012.11.002

Gelin, M., Bonin, P., Méot, A., & Bugaiska, A. (2018). Do animacy effects persist in memory for context? *Quarterly Journal of Experimental Psychology*, *71*(4), 965–974. https://doi.org/10.1080/17470218.2017.1307866

Gerver, C. R., Overman, A. A., Babu, H. J., Hultman, C. E., & Dennis, N. A. (2020). Examining the Neural Basis of Congruent and Incongruent Configural Contexts during Associative Retrieval. *Journal of Cognitive Neuroscience*, 1–17. https://doi.org/10.1162/jocn_a_01593

Ghosh, V. E., & Gilboa, A. (2014). What is a memory schema? A historical perspective on current neuroscience literature. *Neuropsychologia*, *53*, 104–114. https://doi.org/10.1016/j.neuropsychologia.2013.11.010

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology*, *66*(3), 325–331. https://doi.org/10.1111/j.2044-8295.1975.tb01468.x

Greve, A., Cooper, E., Tibon, R., & Henson, R. N. (2019). Knowledge Is Power: Prior Knowledge Aids Memory for Both Congruent and Incongruent Events, but in Different Ways. *Journal of Experimental Psychology: General*, *148*(2), 325–341. http://dx.doi.org/10.1037/xge0000498

Griffiths, B. J., Martín-Buro, M. C., Staresina, B. P., & Hanslmayr, S. (2021). Disentangling neocortical alpha/beta and hippocampal theta/gamma oscillations in human episodic memory formation. *NeuroImage*, *242*, 118454. https://doi.org/10.1016/j.neuroimage.2021.118454

Griffiths, B. J., Parish, G., Roux, F., Michelmann, S., van der Plas, M., Kolibius, L. D., Chelvarajah, R., Rollings, D. T., Sawlani, V., Hamer, H., Gollwitzer, S., Kreiselmeyer, G., Staresina, B., Wimber, M., & Hanslmayr, S. (2019). Directional coupling of slow and fast hippocampal gamma with neocortical alpha/beta oscillations in human episodic memory. *Proceedings of the National Academy of Sciences*, *116*(43), 21834–21842. https://doi.org/10.1073/pnas.1914180116

Haeuser, K. I., & Kray, J. (2023). Effects of prediction error on episodic memory retrieval: Evidence from sentence reading and word recognition. *Language, Cognition and Neuroscience*, *38*(4), 558–574. https://doi.org/10.1080/23273798.2021.1924387

Hamm, J. P., Johnson, B. W., & Kirk, I. J. (2002). Comparison of the N300 and N400 ERPs to picture stimuli in congruent and incongruent contexts. *Clinical Neurophysiology*, *113*(8), 1339–1350. https://doi.org/10.1016/S1388-2457(02)00161-X

Hanczakowski, M., Zawadzka, K., & Macken, B. (2015). Continued effects of context reinstatement in recognition. *Memory & Cognition*, *43*(5), 788–797. https://doi.org/10.3758/s13421-014-0502-2

Hanslmayr, S., Spitzer, B., & Bauml, K.-H. (2009). Brain Oscillations Dissociate between Semantic and Nonsemantic Encoding of Episodic Memories. *Cerebral Cortex*, *19*(7), 1631–1640. https://doi.org/10.1093/cercor/bhn197

Hanslmayr, S., Staresina, B. P., & Bowman, H. (2016). Oscillations and Episodic Memory: Addressing the Synchronization/Desynchronization Conundrum. *Trends in Neurosciences,* *39*(1), 16–25. https://doi.org/10.1016/j.tins.2015.11.004

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

Hayes, S. M., Nadel, L., & Ryan, L. (2007). The effect of scene context on episodic object recognition: Parahippocampal cortex mediates memory encoding and retrieval success. *Hippocampus*, *17*(9), 873–889. https://doi.org/10.1002/hipo.20319

Heib, D. P. J., Hoedlmoser, K., Anderer, P., Gruber, G., Zeitlhofer, J., & Schabus, M. (2015). Oscillatory Theta Activity during Memory Formation and Its Impact on Overnight Consolidation: A Missing Link? *Journal of Cognitive Neuroscience*, *27*(8), 1648–1658. https://doi.org/10.1162/jocn_a_00804

Hollingworth, A. (2006). Scene and Position Specificity in Visual Memory for Objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(1), 58–69. https://doi.org/10.1037/0278-7393.32.1.58

Hollingworth, A. (2007). Object-position binding in visual memory for natural scenes and object arrays. *Journal of Experimental Psychology: Human Perception and Performance,* *33*(1), 31–47. https://doi.org/10.1037/0096-1523.33.1.31

Hollingworth, A., & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental*

Psychology: Human Perception and Performance, *28*(1), 113–136. https://doi.org/10.1037/0096-1523.28.1.113

Höltje, G., Lubahn, B., & Mecklinger, A. (2019). The congruent, the incongruent, and the unexpected: Event-related potentials unveil the processes involved in schematic encoding. *Neuropsychologia*, *131*, 285–293. https://doi.org/10.1016/j.neuropsychologia.2019.05.013

Horner, A. J., Bisby, J. A., Bush, D., Lin, W.-J., & Burgess, N. (2015). Evidence for holistic episodic recollection via hippocampal pattern completion. *Nature Communications*, *6*(1), 7462. https://doi.org/10.1038/ncomms8462

Hovhannisyan, M., Clarke, A., Geib, B. R., Cicchinelli, R., Monge, Z., Worth, T., Szymanski, A., Cabeza, R., & Davis, S. W. (2021). The visual and semantic features that predict object memory: Concept property norms for 1,000 object images. *Memory & Cognition*, *49*(4), 712–731. https://doi.org/10.3758/s13421-020-01130-5

Hupbach, A., Gomez, R., Hardt, O., & Nadel, L. (2007). Reconsolidation of episodic memories: A subtle reminder triggers integration of new information. *Learning & Memory*, *14*(1–2), 47–53. https://doi.org/10.1101/lm.365707

Iber, C., Ancoli-Israel, S., Chesson, A., & Quan, S. F. (2007). AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Techincal Specification. *American Academy of Sleep Medicine*.

Imai, S., & Richman, C. L. (1991). Is the bizarreness effect a special case of sentence reorganization? *Bulletin of the Psychonomic Society*, *29*(5), 429–432. https://doi.org/10.3758/BF03333962

Johns, M. W. (1991). A New Method for Measuring Daytime Sleepiness: The Epworth Sleepiness Scale. *Sleep*, *14*(6), 540–545. https://doi.org/10.1093/sleep/14.6.540

Jones, M. N., Johns, B. T., & Recchia, G. (2012). The role of semantic diversity in lexical organization. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *66*(2), 115–124. https://doi.org/10.1037/a0026727

Jurewicz, K., Cordi, M. J., Staudigl, T., & Rasch, B. (2016). No Evidence for Memory Decontextualization across One Night of Sleep. *Frontiers in Human Neuroscience*, *10*. https://doi.org/10.3389/fnhum.2016.00007

Kleiner, M., Brainard, D., Ingling, A., Murray, R., & Broussard, C. (2007). What’s new in psychtoolbox-3. *Perception*, *36*(14), 1–16.

Klimesch, W., Schimke, H., & Schwaiger, J. (1994). Episodic and semantic memory: An analysis in the EEG theta and alpha band. *Electroencephalography and Clinical Neurophysiology*, *91*(6), 428–441. https://doi.org/10.1016/0013-4694(94)90164-3

Klinzing, J. G., Niethard, N., & Born, J. (2019). Mechanisms of systems memory consolidation during sleep. *Nature Neuroscience*, *22*(10), 1598–1610. https://doi.org/10.1038/s41593-019-0467-3

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. *Journal of Experimental Psychology: General*, *139*(3), 558–578. https://doi.org/10.1037/a0019165

Köster, M., Finger, H., Graetz, S., Kater, M., & Gruber, T. (2018). Theta-gamma coupling binds visual perceptual features in an associative memory task. *Scientific Reports*, *8*(1), 17688. https://doi.org/10.1038/s41598-018-35812-7

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*(12), 463–470. https://doi.org/10.1016/S1364-6613(00)01560-6

Kutas, M., & Hillyard, S. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*(4427), 203–205. https://doi.org/10.1126/science.7350657

Lahl, O., Wispel, C., Willigens, B., & Pietrowsky, R. (2008). An ultra short episode of sleep is sufficient to promote declarative memory performance. *Journal of Sleep Research*, *17*(1), 3–10. https://doi.org/10.1111/j.1365-2869.2008.00622.x

Lampinen, J. M., Copeland, S. M., & Neuschatz, J. S. (2001). Recollections of things schematic: Room schemas revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(5), 1211. https://doi.org/10.1037/0278-7393.27.5.1211

Lau, H., Alger, S. E., & Fishbein, W. (2011). Relational Memory: A Daytime Nap Facilitates the Abstraction of General Concepts. *PLoS ONE*, *6*(11). https://doi.org/10.1371/journal.pone.0027139

Laurent, X., Estévez, A. F., & Marí-Beffa, P. (2020). Evidences for better recall of congruent items in episodic memory. *Cognitive Processing*, *21*(3), 469–477. https://doi.org/10.1007/s10339-020-00963-x

Levi, D. M. (2008). Crowding—An essential bottleneck for object recognition: A mini-review. *Vision Research*, *48*(5), 635–654. https://doi.org/10.1016/j.visres.2007.12.009

MacDonald, K. J., & Cote, K. A. (2021). Contributions of Post-learning REM and NREM Sleep to Memory Retrieval. *Sleep Medicine Reviews*, 101453. https://doi.org/10.1016/j.smrv.2021.101453

Maguire, M. J., & Abel, A. D. (2013). What changes in neural oscillations can reveal about developmental cognitive neuroscience: Language development as a case in point. *Developmental Cognitive Neuroscience*, *6*, 125–136. https://doi.org/10.1016/j.dcn.2013.08.002

Mak, M. H. C., Hsiao, Y., & Nation, K. (2021). Anchoring and contextual variation in the early stages of incidental word learning during reading. *Journal of Memory and Language*, *118*. https://doi.org/10.1016/j.jml.2020.104203

Mak, M. H. C., & Twitchell, H. (2020). Evidence for preferential attachment: Words that are more well connected in semantic networks are better at acquiring new links in paired-associate learning. *Psychonomic Bulletin & Review*, *27*(5), 1059–1069. https://doi.org/10.3758/s13423-020-01773-0

Mander, B. A., Santhanam, S., Saletin, J. M., & Walker, M. P. (2011). Wake deterioration and sleep restoration of human learning. *Current Biology*, *21*(5), R183-R184. https://doi.org/10.1016/j.cub.2011.01.019

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

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Maxcey, A. M., Glenn, H., & Stansberry, E. (2018). Recognition-induced forgetting does not occur for temporally grouped objects unless they are semantically related. *Psychonomic Bulletin & Review; New York*, *25*(3), 1087–1103. https://doi.org/10.3758/s13423-017-1302-z

Mercer, T., & Jones, G. A. (2019). Time-dependent forgetting and retrieval practice effects in detailed visual long-term memory. *Quarterly Journal of Experimental Psychology*, *72*(6), 1561–1577. https://doi.org/10.1177/1747021818799697

Morales, S., & Bowers, M. E. (2022). Time-frequency analysis methods and their application in developmental EEG data. *Developmental Cognitive Neuroscience*, *54*, 101067. https://doi.org/10.1016/j.dcn.2022.101067

Murnane, K., Phelps, M. P., & Malmberg, K. (1999). Context-dependent recognition memory: The ICE theory. *Journal of Experimental Psychology: General*, *128*(4), 403–415. https://doi.org/10.1037/0096-3445.128.4.403

Nadel, L., Campbell, J., & Ryan, L. (2007). Autobiographical Memory Retrieval and Hippocampal Activation as a Function of Repetition and the Passage of Time. *Neural Plasticity*, *2007*, 1–14. https://doi.org/10.1155/2007/90472

Nadel, L., Hupbach, A., Gomez, R., & Newman-Smith, K. (2012). Memory formation, consolidation and transformation. *Neuroscience & Biobehavioral Reviews*, *36*(7), 1640–1645. https://doi.org/10.1016/j.neubiorev.2012.03.001

Nadel, L., & Moscovitch, M. (1997). Memory consolidation, retrograde amnesia and the hippocampal complex. *Current Opinion in Neurobiology*, *7*(2), 217–227. https://doi.org/10.1016/S0959-4388(97)80010-4

Nairne, J. S., VanArsdall, J. E., & Cogdill, M. (2017). Remembering the Living: Episodic Memory Is Tuned to Animacy. *Current Directions in Psychological Science*, *26*(1), 22–27. https://doi.org/10.1177/0963721416667711

Nairne, J. S., VanArsdall, J. E., Pandeirada, J. N. S., Cogdill, M., & LeBreton, J. M. (2013). Adaptive Memory: The Mnemonic Value of Animacy. *Psychological Science*, *24*(10), 2099–2105. https://doi.org/10.1177/0956797613480803

Nakashima, R., & Yokosawa, K. (2011). Does scene context always facilitate retrieval of visual object representations? *Psychonomic Bulletin & Review*, *18*(2), 309–315. https://doi.org/10.3758/s13423-010-0045-x

Neunuebel, J. P., & Knierim, J. J. (2014). CA3 Retrieves Coherent Representations from Degraded Input: Direct Evidence for CA3 Pattern Completion and Dentate Gyrus Pattern Separation. *Neuron*, *81*(2), 416–427. https://doi.org/10.1016/j.neuron.2013.11.017

New, J., Cosmides, L., & Tooby, J. (2007). Category-specific attention for animals reflects ancestral priorities, not expertise. *Proceedings of the National Academy of Sciences*, *104*(42), 16598–16603. https://doi.org/10.1073/pnas.0703913104

Norman, K. A., Detre, G., & Polyn, S. M. (2008). Computational Models of Episodic Memory. In *R. Sun (Ed.), The Cambridge handbook of computational psychology* (pp. 189–225). Cambridge University Press. https://doi.org/10.1017/CBO9780511816772.011

Nuxoll, A., & Laird, J. E. (2004). A cognitive model of episodic memory integrated with a general cognitive architecture. *Lawrence Earlbaum*, 220–225.

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Sciences*, *11*(12), 520–527. https://doi.org/10.1016/j.tics.2007.09.009

O'Neill, J., Pleydell-Bouverie, B., Dupret, D., & Csicsvari, J. (2010). Play it again: Reactivation of waking experience and memory. *Trends in Neurosciences*, *33*(5), 220–229. https://doi.org/10.1016/j.tins.2010.01.006

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*, *2011*, 1–9. https://doi.org/10.1155/2011/156869

Osipova, D., Takashima, A., Oostenveld, R., Fernandez, G., Maris, E., & Jensen, O. (2006). Theta and Gamma Oscillations Predict Encoding and Retrieval of Declarative Memory. *Journal of Neuroscience*, *26*(28), 7523–7531. https://doi.org/10.1523/JNEUROSCI.1948-06.2006

Packard, P. A., Rodríguez-Fornells, A., Bunzeck, X. N., Nicolas, X. B., de Diego-Balaguer, X. R., & Fuentemilla, L. (2017). Semantic Congruence Accelerates the Onset of the Neural Signals of Successful Memory Encoding. *The Journal of Neuroscience*, *37*(2), 291–301. https://doi.org/10.1523/JNEUROSCI.1622-16.2016

Packard, P. A., Steiger, T. K., Fuentemilla, L., & Bunzeck, N. (2020). Neural oscillations and event-related potentials reveal how semantic congruence drives long-term memory in both young and older humans. *Scientific Reports*, *10*(1), 9116. https://doi.org/10.1038/s41598-020-65872-7

Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition*, *3*(5), 519–526. https://doi.org/10.3758/BF03197524

Pelli, D. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442. https://doi.org/10.1163/156856897X00366

Pezdek, K., Maki, R., Valencia-Laver, D., Whetstone, T., Stoeckert, J., & Dougherty, T. (1988). Picture memory: Recognizing added and deleted details. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(3), 468. https://doi.org/10.1037/0278-7393.14.3.468

Pezdek, K., Whetstone, T., Reynolds, K., Askari, N., & Dougherty, T. (1989). Memory for Real-World Scenes: The Role of Consistency With Schema Expectation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(4), 587–595. https://doi.org/10.1037/0278-7393.15.4.587

Piolino, P., Desgranges, B., & Eustache, F. (2009). Episodic autobiographical memories over the course of time: Cognitive, neuropsychological and neuroimaging findings. *Neuropsychologia*, *47*(11), 2314–2329. https://doi.org/10.1016/j.neuropsychologia.2009.01.020

Potter, M. C., Staub, A., Rado, J., & O’Connor, D. H. (2002). Recognition memory for briefly presented pictures: The time course of rapid forgetting. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(5), 1163. https://doi.org/10.1037/0096-1523.28.5.1163

Rauchs, G., Desgranges, B., Foret, J., & Eustache, F. (2005). The relationships between memory systems and sleep stages. *Journal of Sleep Research*, *14*(2), 123–140. https://doi.org/10.1111/j.1365-2869.2005.00450.x

Reggev, N., Sharoni, R., & Maril, A. (2018). Distinctiveness Benefits Novelty (and Not Familiarity), but Only Up to a Limit: The Prior Knowledge

Perspective. *Cognitive Science*, *42*(1), 103–128. https://doi.org/10.1111/cogs.12498

Reiner, M., Rozengurt, R., & Barnea, A. (2014). Better than sleep: Theta neurofeedback training accelerates memory consolidation. *Biological Psychology*, *95*, 45–53. https://doi.org/10.1016/j.biopsycho.2013.10.010

Reinitz, M. T., Peria, W. J., Séguin, J. A., & Loftus, G. R. (2011). Different confidence–accuracy relationships for feature-based and familiarity-based memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(2), 507–515. https://doi.org/10.1037/a0021961

Reinitz, M. T., Séguin, J. A., Peria, W., & Loftus, G. R. (2012). Confidence–accuracy relations for faces and scenes: Roles of features and familiarity. *Psychonomic Bulletin & Review*, *19*(6), 1085–1093. https://doi.org/10.3758/s13423-012-0308-9

Roesler, R., & McGaugh, J. L. (2019). Memory Consolidation. In *Reference Module in Neuroscience and Biobehavioral Psychology*. Elsevier. https://doi.org/10.1016/B978-0-12-809324-5.21493-4

Roux, F., Parish, G., Chelvarajah, R., Rollings, D. T., Sawlani, V., Hamer, H., Gollwitzer, S., Kreiselmeyer, G., Ter Wal, M. J., Kolibius, L., Staresina, B. P., Wimber, M., Self, M. W., & Hanslmayr, S. (2022). Oscillations support short latency co-firing of neurons during human episodic memory formation. *eLife*, *11*, e78109. https://doi.org/10.7554/eLife.78109

Schneider, J. M., & Maguire, M. J. (2018). Identifying the relationship between oscillatory dynamics and event-related responses. *International Journal of Psychophysiology*, *133*, 182–192. https://doi.org/10.1016/j.ijpsycho.2018.07.002

Schreiner, M. R., Meiser, T., & Bröder, A. (2023). The binding structure of event elements in episodic memory and the role of animacy. *Quarterly Journal of Experimental Psychology*, *76*(4), 705–730. https://doi.org/10.1177/17470218221096148

Schreiner, T., & Staudigl, T. (2020). Electrophysiological signatures of memory reactivation in humans. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1799), 20190293. https://doi.org/10.1098/rstb.2019.0293

Shahid, A., Wilkinson, K., Marcu, S., & Shapiro, C. M. (2011). Stanford Sleepiness Scale (SSS). In A. Shahid, K. Wilkinson, S. Marcu, & C. M. Shapiro (Eds.), *STOP, THAT and One Hundred Other Sleep Scales* (pp. 369–370). Springer New York. https://doi.org/10.1007/978-1-4419-9893-4_91

Shepard, R. N. (1967). Recognition Memory for Words, Sentences, and Pictures. *Journal of Verbal Learning and Verbal Behavior*, *6*(1), 156–163. https://doi.org/10.1016/S0022-5371(67)80067-7

Shoben, E. J., Wescourt, K. T., & Smith, E. E. (1978). Sentence Verification, Sentence Recognition, and the Semantic-Episodic Distinction. *Journal of Experimental Psychology: Human Learning and Memory*, *4*(4), 304–317. https://doi.org/10.1037/0278-7393.4.4.304

Smith, S. M., Handy, J. D., Hernandez, A., & Jacoby, L. L. (2018). Context specificity of automatic influences of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(10), 1501–1513. https://doi.org/10.1037/xlm0000523

Squire, L., & Zola-Morgan, S. (1991). The medial temporal lobe memory system. *Science*, *253*(5026), 1380–1386. https://doi.org/10.1126/science.1896849

Standing, L. (1973). Learning 10000 pictures. *Quarterly Journal of Experimental Psychology*, *25*(2), 207–222. https://doi.org/10.1080/14640747308400340

Standing, L., Conezio, J., & Haber, R. N. (1970). Perception and memory for pictures: Single-trail learning of 2500 visual stimuli. *Psychonomic Science*, *19*, 73–74.

Staresina, B. P., Gray, J. C., & Davachi, L. (2009). Event Congruency Enhances Episodic Memory Encoding through Semantic Elaboration and Relational Binding. *Cerebral Cortex*, *19*(5), 1198–1207. https://doi.org/10.1093/cercor/bhn165

Staudigl, T., & Hanslmayr, S. (2013). Theta Oscillations at Encoding Mediate the Context-Dependent Nature of Human Episodic Memory. *Current Biology*, *23*(12), 1101–1106. https://doi.org/10.1016/j.cub.2013.04.074

Sun, H.-M., Simon-Dack, S. L., Gordon, R. D., & Teder, W. A. (2011). Contextual influences on rapid object categorization in natural scenes.

Brain Research, *1398*, 40–54. https://doi.org/10.1016/j.brainres.2011.04.029

Sutherland, G. (2000). Memory trace reactivation in hippocampal and neocortical neuronal ensembles. *Current Opinion in Neurobiology*, *10*(2), 180–186. https://doi.org/10.1016/S0959-4388(00)00079-9

Sweegers, C. C. G., & Talamini, L. M. (2014). Generalization from episodic memories across time: A route for semantic knowledge acquisition. *Cortex*, *59*, 49–61. https://doi.org/10.1016/j.cortex.2014.07.006

Tambini, A., & D'Esposito, M. (2020). Causal Contribution of Awake Post-encoding Processes to Episodic Memory Consolidation. *Current Biology*, *30*(18), 3533-3543.e7. https://doi.org/10.1016/j.cub.2020.06.063

Thornhill, D. E., & Van Petten, C. (2012). Lexical versus conceptual anticipation during sentence processing: Frontal positivity and N400 ERP components. *International Journal of Psychophysiology*, *83*(3), 382–392. https://doi.org/10.1016/j.ijpsycho.2011.12.007

Tibon, R., Cooper, E., & Greve, A. (2017). Does Semantic Congruency Accelerate Episodic Encoding, or Increase Semantic Elaboration? *The Journal of Neuroscience*, *37*(19), 4861–4863. https://doi.org/10.1523/JNEUROSCI.0570-17.2017

Tononi, G., & Cirelli, C. (2003). Sleep and synaptic homeostasis: A hypothesis. *Brain Research Bulletin*, *62*(2), 143–150. https://doi.org/10.1016/j.brainresbull.2003.09.004

Tononi, G., & Cirelli, C. (2006). Sleep function and synaptic homeostasis. *Sleep Medicine Reviews*, *10*(1), 49–62. https://doi.org/10.1016/j.smrv.2005.05.002

Tse, D., Langston, R. F., Kakeyama, M., Bethus, I., Spooner, P. A., Wood, E. R., Witter, M. P., & Morris, R. G. M. (2007). Schemas and Memory Consolidation. *Science*, *316*(5821), 76–82. https://doi.org/10.1126/science.1135935

Tucker, M. A., & Fishbein, W. (2008). Enhancement of Declarative Memory Performance Following a Daytime Nap Is Contingent on Strength of Initial Task Acquisition. *Sleep*, *31*(2), 197-203. https://dx.doi.org/10.1093%2Fsleep%2F31.2.197

Tulving, E. (1983). *Elements of episodic memory*. Oxford: Clarendon.

Tulving, E., & Thomas, D. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*, 352.

van Kesteren, M. T. R., Rignanese, P., Gianferrara, P. G., Krabbendam, L., & Meeter, M. (2020). Congruency and reactivation aid memory integration through reinstatement of prior knowledge. *Scientific Reports*, *10*(1), 4776. https://doi.org/10.1038/s41598-020-61737-1

VanArsdall, J. E., Nairne, J. S., Pandeirada, J. N. S., & Blunt, J. R. (2013). Adaptive Memory: Animacy Processing Produces Mnemonic Advantages. *Experimental Psychology*, *60*(3), 172–178. https://doi.org/10.1027/1618-3169/a000186

Vogt, S., & Magnussen, S. (2007). Long-Term Memory for 400 Pictures on a Common Theme. *Experimental Psychology*, *54*(4), 298–303. https://doi.org/10.1027/1618-3169.54.4.298

Wamsley, E. J. (2019). Memory Consolidation during Waking Rest. *Trends in Cognitive Sciences*, *23*(3), 171–173. https://doi.org/10.1016/j.tics.2018.12.007

Wang, W., Subagdja, B., Tan, A., & Starzyk, J. A. (2012). Neural Modeling of Episodic Memory: Encoding, Retrieval, and Forgetting. *IEEE Transactions on Neural Networks and Learning Systems*, *23*(10), 1574–1586. IEEE Transactions on Neural Networks and Learning Systems. https://doi.org/10.1109/TNNLS.2012.2208477

Wiltgen, B. J., Zhou, M., Cai, Y., Balaji, J., Karlsson, M. G., Parivash, S. N., Li, W., & Silva, A. J. (2010). The Hippocampus Plays a Selective Role in the Retrieval of Detailed Contextual Memories. *Current Biology*, *20*(15), 1336–1344. https://doi.org/10.1016/j.cub.2010.06.068

Winocur, G., Frankland, P. W., Sekeres, M., Fogel, S., & Moscovitch, M. (2009). Changes in context-specificity during memory reconsolidation: Selective effects of hippocampal lesions. *Learning & Memory*, *16*(11), 722–729. https://doi.org/10.1101/lm.1447209

Winocur, G., & Moscovitch, M. (2011). Memory Transformation and Systems Consolidation. *Journal of the International Neuropsychological Society*, *17*(05), 766–780. https://doi.org/10.1017/S1355617711000683

Winocur, G., Moscovitch, M., & Sekeres, M. (2007). Memory consolidation or transformation: Context manipulation and hippocampal representations of memory. *Nature Neuroscience*, *10*(5), 555–557. https://doi.org/10.1038/nn1880

Woodman, G. F. (2010). A brief introduction to the use of event-related potentials in studies of perception and attention. *Attention, Perception, & Psychophysics*, *72*(8), 2031–2046. https://doi.org/10.3758/BF03196680

Yonelinas, A. P., Ranganath, C., Ekstrom, A. D., & Wiltgen, B. J. (2019). A contextual binding theory of episodic memory: Systems consolidation reconsidered. *Nature Reviews. Neuroscience; London*, *20*(6), 364–375. https://doi.org/10.1038/s41583-019-0150-4

Zhang, J. (2022). The role of sleep for episodic memory consolidation: Stabilizing or rescuing? *Neurobiology of Learning and Memory*, 6.

Zhao, C., Fukuda, K., & Woodman, G. F. (2022). Cross-frequency coupling of frontal theta and posterior alpha is unrelated to the fidelity of visual long-term memory encoding. *Visual Cognition*, 1–14. https://doi.org/10.1080/13506285.2022.2084480

Zhu, Z., Wang, Y., Jia, J., & Wu, Y. (2019). Beta and alpha desynchronizations underlie reconsolidation-mediated episodic memory updating. *Neuropsychologia*, *132*, 107135. https://doi.org/10.1016/j.neuropsychologia.2019.107135

Zöllner, C., Klein, N., Cheng, S., Schubotz, R. I., Axmacher, N., & Wolf, O. T. (2023). Where was the toaster? A systematic investigation of semantic construction in a new virtual episodic memory paradigm. *Quarterly Journal of Experimental Psychology*, *76*(7), 1497–1514. https://doi.org/10.1177/17470218221116610