Predictive Processing and Schema Effects in Episodic Memory

Adam J. Curtis

PhD

University of York

Psychology

September 2025

Abstract

Schemas are associative knowledge structures that represent statistical regularities across multiple episodes. Schemas are known to enhance memory retention, with better memory observed for events that are either highly congruent or highly incongruent with our existing schemas. These schema effects are thought to arise due to predictive processing, whereby prior knowledge is used to make predictions about how an event will be structured. The subsequent affirmation or violation of these predictions is believed to potentiate memory encoding via prefrontal-hippocampal interactions. This thesis investigates whether schema effects extend to situations in which a core feature of the schema is the chronology of events, allowing for temporal predictions of the future. Participants were pre-trained on categorical sequences where schemas defined the order of semantic categories. Subsequently, they studied new sequences consisting of novel exemplars. Crucially, these sequences could either be congruent, incongruent, or unrelated to the pre-learned schemas. At test, participants were cued with an item from a sequence and required to recall the following two items in a four-alternative forced-choice task. Schema congruency effects were observed in all experiments. However, incongruency benefits were only observed for categorical violations, and not for violations of sequential order. MEG analyses revealed that schema violations were associated with modulations of theta-band activity, and that category, sequence position, and uncertainty signals could all be decoded from pre-stimulus neural activity. Overall, these findings are in line with predictive accounts of schema effects and refine our understanding of when memory advantages arise for schematically incongruent sequences.

Table of contents

List of Tables

List of Figures

Acknowledgements

Personal

I would like to thank my Father, Mark Gerard Curtis. Without his ongoing support over my past decade as a student, this thesis would not have been possible.

And my partner, Amber Bonser, for always being there without having to be asked.

Academic

I would like to thank the following people for their guidance and feedback throughout my PhD. Their ongoing support has been invaluable to my growth as a scientist.

- Professor Aidan J. Horner (PhD Supervisor)
- Professor Maria Wimber (Collaborator)
- Professor Alex Wade and Professor Beth Jefferies (Thesis Advisory Panel)
- All members of York Episodic Memory lab: Chloe Brunskill, Jamie Cockcroft, Darcy Denneen, Arianna Moccia, Ewan Murray, Sophie O’Reilly, Aida Rincon Chong, and Akul Satish.
- All staff at York Neuroimaging Centre (YNiC)

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for a degree or other qualification at this University or elsewhere. All sources are acknowledged as references. I acknowledge that I received assistance from chatGPT to proofread this thesis in line with the *Policy on Transparency in Authorship in PGR Programmes.*

1 Literature Review

1.1 Introduction

Memory exerts a pervasive influence on perception and cognition. Indeed, the ability to remember the past is the foundation of intelligent behaviour. Without an internal model of how the world tends to work, flexible and adaptive responses to novel situations would be severely constrained. Every time we experience an event, the sights, sounds, and emotions are all interpreted with regard to the vast body of knowledge we already possess. Just as perception and action guide what we remember, what we remember guides what we perceive and do.

One way in which this generalisation of knowledge across time is achieved is via the construction and application of mnemonic schemas. Schemas are associative knowledge structures that represent the statistical regularities that occur across multiple episodes (Ghosh & Gilboa, 2014; Gilboa & Marlatte, 2017). Fundamentally, schemas are mental abstractions that reduce the dimensionality of the knowledge we accrue down to the commonalities that are most predictable. A common example of schema that most people would possess would be a schema for a visit to a restaurant. This schema would comprise an associative network structure with nodes connecting related concepts such as getting a table, reading a menu, ordering and receiving food, and paying the bill. Crucially, the schema would concern the regularities that are common across multiple events, and would lack details specific to any single restaurant visit, for example, the name of the waiter. Over several decades, schemas have featured prominently in memory research as there is robust evidence that they augment learning, enhancing memory retention for information that is either highly congruent or highly incongruent with their content (Brewer & Treyens, 1981; Frank et al., 2018; Greve et al., 2019; Hastie & Kumar, 1979; Ortiz-Tudela et al., 2024; Tse et al., 2007; van Kesteren et al., 2010; Yacoby et al., 2025). These increases in memorability for schema-related content are what are referred to as *mnemonic schema effects*. For brevity, I will refer to them as *schema effects* throughout this thesis.

Schema effects are thought to arise due to *predictive processing*, whereby prior knowledge is used to make predictions about how an event should be structured. The use of prior knowledge to make predictions about the environment, and then to update this knowledge based on the congruence of these predictions with subsequent sensory inputs, may be at the very core of how the human brain functions. Predictive processing theories propose that brain function can be conceptualised as a Bayesian generative model that continuously aims to minimise prediction errors (or, more technically, *free energy*) across time (Friston, 2005; Rao & Ballard, 1999). Such theories provide a potentially unifying account of neurocognitive function that can account for phenomena across multiple cognitive domains including learning and memory, sensory perception, and motor control. By using our prior knowledge of how the world tends to work, the brain can assert top-down influence on our perceptions, actions, and subsequent memory, thereby facilitating adaptive behaviour.

There is a broad consensus that schema effects are mediated by interactions between the medial prefrontal cortex (mPFC) and regions of the medial temporal lobes (MTL) such as the hippocampus (HPC) (Bein et al., 2014; Brod et al., 2016; Gilboa & Marlatte, 2017; Raykov et al., 2019; Robin & Moscovitch, 2017; van Kesteren et al., 2012, 2013). However, it is also thought that these interactions are themselves a consequence of a predictive coding process. Under this model, schemas are thought to support ongoing neural predictions. These predictions are then compared with sensory inputs, which either affirm or violate expectations. This ongoing predictive processing modulates activity in the mPFC-MTL axis, which in turn gives rise to the behavioural manifestation of schema effects in memory.

This thesis is aimed at investigating this predictive account of schema effects in episodic memory. Specifically, it examines whether schema effects can be observed in a temporal order memory paradigm where the chronology of events forms a central component of a schema's content. It then asks whether this knowledge of chronology facilitates temporal predictions of the future, and how these might relate to schema effects in memory. This is done by combining a temporal order memory paradigm with magnetoencephalography (MEG) and multivariate decoding (MVPA).

Before continuing to my experimental work, I will begin with an introduction to the history of schemas in the psychology of memory. I will then discuss the neuroscience of episodic memory more generally, and its relationship to the medial temporal lobes and neuronal oscillations. From there, I will discuss neurocognitive models of PFC-MTL interaction and how these are linked to the influence of schemas on mnemonic processing. This will lead to a discussion of predictive coding theories and the role of prediction errors in learning. Finally, I will discuss sequence learning, the neuroscience of temporal order memory, and how this thesis will contribute to ongoing debates in these fields.

1.2 Schemas and memory

Mnemonic schemas are associative knowledge structures that represent statistical regularities across multiple episodes. Schemas are believed to play an adaptive role in behaviour by facilitating efficient cognitive processing. The notion of a schema in memory research first came to prominence through the work of Frederic Bartlett (Bartlett, 1932). Bartlett's key insight was that our existing schematic knowledge of the world can have a large impact on what we subsequently remember. Bartlett (1932) demonstrated this by presenting western participants with a native American folk tale: The War of the Ghosts. This tale contained supernatural aspects that would be unfamiliar to those who were raised in a western, traditionally European culture. When asked to recall the tale, participants were likely to distort the story to make it more compatible with European cultural norms. Bartlett argued that this indicates a tendency in the participants to use their existing (western) schemas to aid in recalling the story's elements. Although Bartlett's work has been criticised for its lack of formal controls (Gauld & Stephenson, 1967), it nonetheless spawned a wave of research into the influence of schemas on memory.

A key issue throughout the history of schema research has been a lack of consensus regarding the precise definition of a schema. This issue has been most formally addressed by Ghosh & Gilboa (2014) who note the heterogeneous use of the term, the issues that this has created, and some suggestions for how a schema should be formally defined. Ghosh and Gilboa argue that schemas have four defining features: *an associative network structure, a basis on multiple episodes, lack of unit detail,* and *adaptability.* Schemas have an associative network structure in that they are composed of individual units and the connections between them. Crucially, schemas represent commonalities across multiple episodes and hence lack detail regarding the specifics of any single instance. This places them in contrast to episodic memories that are based on a single episode and include specific details about that instance. This basis on multiple episodes and lack of unit detail allows information stored in a schema to be generalised to novel situations and thus allows an individual to use past experiences to behave adaptively. Finally, schemas have adaptability in that they are dynamic representations that can be altered by new experiences. Schemas can either *assimilate* new information, by incorporating new elements without changing existing relationships, or *accommodate* information, by which the fundamental structure of the schema is altered.

In addition to these four defining features, Ghosh & Gilboa (2014) also proposed four features to which schemas are sensitive, but that are not defining features: *chronological* *relationships, hierarchical organisation, cross-connectivity,* and *embedded response options.* Of particular note for the present thesis is a schema's sensitivity to chronology. Knowledge of temporal order may allow an individual to predict how an event is likely to unfold, the neural basis of which remains largely unexplored in schema memory research. Given the ambiguous use of the term *schema* throughout the memory literature, the present thesis aims to adhere as much as possible to the definition put forward by Ghosh & Gilboa (2014).

Research suggests that experiences that are congruent with our existing schemas tend to be better remembered than experiences that are not. One of the earliest demonstrations of this effect investigated memory for real-life locations (Brewer & Treyens, 1981). In this experiment, participants were asked to wait in a room that was arranged to resemble a graduate student's office before engaging in a surprise memory test. Memory was tested using free recall and recognition. It was found that items that were congruent with an office schema (e.g., desk, chair, typewriter) were better remembered than items that were schema-inconsistent (e.g., skull, brick, picnic basket). These findings were interpreted as being a consequence of the participants' long-term schematic knowledge of the typical structure of an office environment, highlighting an important role for previous knowledge in shaping new memories.

Since this initial finding, congruency advantages have been observed in paradigms using stimuli as diverse as colour-item pairings (Staresina et al., 2008), item-location pairings (Buuren et al., 2014), word lists (Bein et al., 2015), films (van Kesteren et al., 2010), stories (Mandler & Johnson, 1977), and, in rats, flavour-place associations (Tse et al., 2007). For example, in an fMRI study, Staresina et al. (2008) presented participants with colour-item pairings that varied in their plausibility (e.g., red-elephant). Participants were asked to either give a plausibility rating (how likely the pair was to occur in real life) or a valence rating (how appealing or unappealing the pair was). Colour-item pairings would be classified as schema-congruent if the participant judged them as plausible or appealing.

During a surprise memory test, participants performed an old-new recognition test on the words. For words that had been judged as old, participants were asked to retrieve the colour that was associated with the word and the task (i.e., plausibility or valence) that it appeared in. A congruency effect was observed for both word recognition and colour recall with better memory for colour-item pairings that were judged as plausible or appealing in the encoding phase. Importantly, a congruency effect was also observed for recall of the task each colour-item pair appeared in. This is important because it suggests that the congruency advantage extends not only to the items under evaluation but to the entire encoding context. This is but one example, however, the finding that schema-congruent information is better remembered than information that is unrelated to such knowledge appears to be a robust phenomenon.

Somewhat paradoxically, other strands of research suggest that surprise (i.e., events that do not conform to our prior experiences) also enhances learning. Indeed, the notion of a prediction error features prominently in many mathematical models of learning (Pearce & Hall, 1980; Rescorla & Wagner, 1972). Early work examining this idea investigated memory for behaviours that were either congruent, incongruent or neutral with regard to different personality traits (Hastie & Kumar, 1979). Participants were presented with trait adjectives describing a character and then read 20 sentences describing behaviours performed by that character. Memory for these behaviours was then tested using a free recall task. Behaviours that contradicted the trait impressions were recalled more frequently than congruent or neutral behaviours, suggesting enhanced memory for schema-incongruent events.

In a more recent study, Kafkas & Montaldi (2018) had participants learn cue-item contingencies in which a cue predicted whether a following item would be natural or man-made. Expectation violations were then manipulated to occur during either encoding or retrieval. In the encoding violation experiments, participants first learned the cue-item contingencies before encoding cue-item pairings that either did or did not conform to the contingencies.

At test, participants performed a recognition task on the items, judging them as either familiar, recollected, or new. In the retrieval violation experiments, participants first encoded the items by making either a size judgement or a man-made vs. natural judgement before learning the cue-item contingencies. Expectation violations were then generated during the recognition task by preceding each item with a symbol that either did or did not conform to the pre-learned contingencies. It was found that recollection memory was enhanced for unexpected items regardless of whether the expectation violation occurred at encoding or retrieval. In addition to this finding, similar incongruency advantages have been observed for scene-item pairings (Greve et al., 2017), object-environment contingencies (Mäntylä & Bäckman, 1992), and word lists (Tulving & Kroll, 1995). In all, this prompts the question as to why we observe such a division regarding whether schema-congruent or schema-incongruent information most benefits memory retention.

A potential answer to this conflict has been presented by Greve et al. (2019) who were testing the behavioural predictions of the SLIMM (schema-linked interactions between medial prefrontal and medial temporal regions) framework (van Kesteren et al., 2012). SLIMM is a neurocognitive framework that attempts to account for schema effects at a systems neuroscience level. I will return to SLIMM in more detail in the section on models of PFC-MTL interaction. Greve et al. (2019) proposed that the seemingly contradictory results in the schema literature may be due to researchers treating congruency as a binary variable (i.e., congruent vs. incongruent), and then biasing their studies to heavily focus on one side of this dichotomy. Instead, schema congruency may be more accurately conceptualised as a spectrum, with at least three levels of schema congruency being required to accurately represent how new information relates to a schema. More specifically, it is possible for new information to be either weakly or strongly incongruent with a schema. Weakly incongruent information may simply be *unrelated* to the schema. Without actively contradicting the schema's content, such a relationship would be unlikely to generate a strong prediction error and thus would lead to little learning. On the other hand, strongly incongruent information may actively contradict a schema's content, generating a strong prediction error and thus potentiating learning. Under this framework, memory retention is predicted to be a U-shaped function of schema congruency.

Greve et al. (2019) provided evidence for this U-shaped function of schema congruency. During study, participants learned which of two items (e.g., shoe, umbrella, chair) had the highest value via feedback. Participants were presented with multiples of each item (e.g., two umbrellas, three shoes) and had to infer through trial-and-error which had the highest value after being told that one item had twice the value of the other. Pairs of items could appear in one of three conditions: congruent, in which the rule remained the same on all study trials, unrelated, in which there was no consistent rule throughout study, and incongruent, in which there was initially a consistent rule, but this rule reversed part way through study. At test, participants performed an old-new recognition task on the displays that were viewed during study trials. It was found that old-new recognition performance was a U-shaped function of schema congruency with displays from the congruent and incongruent conditions being better remembered than displays from the unrelated condition. More recently, such a U-shaped function has also been demonstrated in an object-location virtual reality paradigm in which participants recalled the locations of objects which varied in their congruency with a kitchen schema (Quent et al., 2022). Interestingly, the presence of this U-shaped function also seems to be age-dependent, with the benefits of schema incongruency being much more prevalent in younger rather than older adults (Ortiz-Tudela et al., 2024).

Most recently, studies have revealed that the effects of prior knowledge on memory retention may be dependent on the retrieval goals and orientation of the task. Yacoby et al. (2025) attempted to test the predictions of the SLIMM framework when they were applied to different memory tests and task goals. Across six experiments, participants were tested on their item memory (old vs. new), congruency memory (whether an item occurred in an expected or unexpected context), and/or item-context associative memory (two- or four-alternative forced-choice) either immediately after study or after a 24-hour delay. During study, participants would see a scene (e.g., a person in a launderette) before an object would appear in the scene (e.g., a doll). While studying the scene-item pairs, participants were required to make an expectancy judgement (how well the item paired with the scene) on a 1-4 scale from highly expected to highly surprising. At test, participants engaged in on old-new recognition task on the items, accompanied by a measure of recollection vs. familiarity. After the recognition test, there was a congruency test, in which participants viewed the items and recalled whether the item occurred in an expected or surprising context (the scene itself did not have to be remembered, just the state of expectedness or surprise). Following the congruency test, participants engaged in an two- or four-alternative forced-choice task in which they retrieved the scene associated with the item. The only metric for which a U-shaped function was observed was congruency memory, however, the presence of this U was reliant on the retrieval orientation of the participant. When participants were queried as to how *surprising* an event was (i.e., how surprised were they by the scene-object pair), the U was present, but when participants were queried on how *expected* an event was (i.e., how much they expected the scene-object pair), only surprising events showed a memory advantage. In the case of item memory and associative memory, only a schema congruency advantage was observed. Moreover, in the case of associative memory, this congruency advantage was likely due to retrieval-based strategies. When these strategies were prevented, the opposite pattern was observed with better memory for unexpected trials. The findings of Yacoby et al. (2025) highlight that the U-shaped curve predicted by the SLIMM can be highly sensitive to task goals and the metrics used to assess memory, suggesting that further theoretical refinement is still required.

One aim of the current thesis is to see if this U-shaped function can be reproduced in a paradigm that builds upon those previously used. Importantly, previous studies have tended to use schemas that are somewhat static, requiring knowledge of contextual rules, spatial organisations, or associative relationships, and have not required temporal predictions of upcoming events. As Ghosh & Gilboa (2014) point out, due to their event-based nature, a significant subset of schemas should be sensitive to chronology and the temporal ordering of events. Indeed, the ability to accurately predict what will happen next in an unfolding sequence is likely a core adaptive advantage of possessing schemas, yet this temporal component is missing from most previous studies. As such, I ask whether the introduction of a temporal component to the schemas, and to the information to be remembered, influences schema effects in memory.

If schemas do enhance memory for sequences of events by facilitating predictions of the future, we would expect this effect to be especially prominent for episodic memories, as episodic memories are defined by their spatiotemporal structure. Next, I therefore move on to discuss episodic memory more generally, how it relates to the medial temporal lobes and hippocampus, and discuss the various neural oscillations that have been associated with episodic memory function.

1.3 Episodic memory, the medial temporal lobes, and neural oscillations

Episodic memory refers to memory for personally experienced events that occurred at specific places and times (Tulving, 2002). At the level of the brain, episodic memory has been most closely associated with the medial temporal lobes (MTL) consisting of the hippocampus (HPC) and the surrounding entorhinal, perirhinal, and parahippocampal cortices. The relationship between the MTL and episodic memory first became apparent in the seminal studies of the neurological patient Henry Molaison (H.M.) (Scoville & Milner, 1957). H.M. contracted severe anterograde amnesia, the loss of ability to form new episodic memories, after having his HPC and other MTL regions resected in an attempt to alleviate his epilepsy. H.M.'s case played a pivotal role in demonstrating that episodic memory is in some ways distinct from other forms of memory such as short-term working memory, semantic memory, and forms of implicit memory such as procedural learning.

Following the seminal case of H.M., further neuropsychological cases emphasised the role of the hippocampus in different domains of memory processing. The patient Y.R. had selective bilateral hippocampal damage. While Y.R. demonstrated relatively normal recognition memory, they showed deficits in the recall of associations, especially associations between two different types of information (e.g., object-location pairs as opposed to object-object pairs) (Mayes et al., 2004). These findings indicated that the hippocampus is critical for *recollection*, where recall includes details of the encoding context in addition to the remembered item itself. This stands in contrast to *familiarity*, that involves a sense of prior encounter with an item that does not include contextual details and which is thought to rely more heavily on structures such as the perirhinal cortex (Bowles et al., 2007; Diana et al., 2007). More recently, research has helped to verify these findings using patients with forms of encephalitis that produce highly selective hippocampal lesions that are homogenous across the population (Lad et al., 2019). Participants in this study demonstrated deficits in episodic but not semantic memory, as well as impaired recollection but preserved recognition. Taken together, these neuropsychological findings provide converging evidence that the hippocampus plays a selective role in episodic recollection. But precisely how does the hippocampus facilitate such recall?

At the cellular level, the hippocampus contains cells that represent stimuli which can be thought of as the subunits, or building blocks, of episodic memories. Most prominent among these are *concept cells*, neurons which respond selectively to abstract concepts such as people and objects (Quiroga et al., 2005; Quiroga, 2012), and *place cells*, neurons which respond selectively to specific locations in an environment (O'Keefe & Dostrovsky, 1971). Moreover, there is also evidence to suggest that the hippocampus contains *time cells*, neurons that code for the sequential ordering of events (Eichenbaum, 2013), and *episode-specific neurons* that code for the conjunction of elements within a single event

(Kolibius et al., 2023). Given the existence of these cell types, the hippocampus may act as an associative hub which binds concepts, places, and temporal order information to form a single event engram. Such a convergence zone provides a region from which the sensory features that constitute an event can be reactivated during recollection (Damasio, 1989).

When we experience an event, different aspects of that experience are represented by the appropriate areas of the neocortex—visual aspects by visual cortex, auditory aspects by auditory cortex, and so on. Recollection is thought to occur when activity in these sensory regions is reinstated at a later date (Wheeler et al., 2000). The hippocampus is believed to act as a mnemonic index storing a trace for these patterns of neocortical activity and triggering their reinstatement during memory retrieval (Horner et al., 2015; Teyler & Rudy, 2007). Under Systems Consolidation Theory (SCT), episodic memories are initially highly reliant on the hippocampus for their maintenance and retrieval. However, over time, and through processes of online rehearsal and offline consolidation, memories are transferred to permanent storage in the neocortex where they become independent of the hippocampus. Despite the prominence of this standard model of systems consolidation, it is not without its critics, and has been challenged by the proposal of Trace Transformation Theory and its predecessor Multiple Trace Theory (Nadel et al., 2007; Nadel & Moscovitch, 1997). Trace Transformation Theory (TTT) proposes that the hippocampus codes not for the remoteness of a memory, as predicted by SCT, but for episodic detail. As such, under TTT, memories only become independent of the hippocampus if semanticisation has occurred, a process by which a memory undergoes an episodic-to-semantic transformation with time. Whether SCT or TTT is a more accurate model of memory consolidation is still a topic of debate. Nevertheless, there is strong evidence to suggest that the hippocampus plays an integral role in the rapid acquisition of instances and the subsequent consolidation of these memories into long-term storage.

The notion of the hippocampus as a mnemonic index has a long history in memory science. However, there is an emerging body of evidence to suggest that the hippocampus also functions as a generative model of the environment that can predict expected sensory inputs based on prior experience. Under such a framework, the hippocampus works as a generative hub that detects prediction errors based on comparisons of sensory experience with expected perceptual inputs (J. Lisman & Redish, 2009; Stachenfeld et al., 2017). At first glance, these differing roles of prediction and reinstatement may seem at odds. Predictive coding theories suggest that higher-order brain regions, such as the hippocampus, should use prediction to inhibit activity regarding expected sensory inputs so as to reduce prediction error. However, mnemonic indexing suggests that the hippocampus should excite, rather than inhibit, the neocortex to initiate recollection. Importantly, recent models have proposed solutions to this apparent contradiction by which both excitation and inhibition of the neocortex can be facilitated by modulating the precision of prediction error units (Barron et al., 2020). I will return to the predictive nature of the hippocampus in the section on predictive processing and memory. For now, I will move on to discuss the neural oscillations that are associated with human episodic memory.

Given that this thesis will be making use of electrophysiological measures, it is worth highlighting the abundance of evidence regarding the importance of neural oscillations in medial temporal lobe and episodic memory function. Theta rhythms (4 - 8 Hz) in particular are known to play a crucial role in medial temporal lobe function. For example, theta oscillations have been linked to spatial navigation (Burgess et al., 2002; Chrastil et al., 2022; Maguire et al., 1999). Hippocampal place cells have been shown to fire at progressively earlier phases of the theta cycle as a rat traverses a cell's firing field, a phenomenon known as *phase precession* that likely relates to a code for the sequencing of locations in space (Buzsáki & Tingley, 2018). In addition to its role in spatial processing, hippocampal theta is also associated with successful encoding and retrieval. For example, intracranial electroencephalography (iEEG) studies have demonstrated that successful associative memory encoding is associated with increases in theta power (Joensen et al., 2023; Kota et al., 2020; Lin et al., 2017). Moreover, the timing of neocortical reactivations during memory retrieval has been linked to hippocampal theta phase, indicating a mechanism by which theta oscillations prevent interference between ongoing encoding and retrieval processes (Kerrén et al., 2018). Interestingly, there is somewhat of a conflict as to whether theta power increases or decreases are most associated with enhanced memory processing. A recent review suggests that this discrepancy may be related to the neuroimaging method (intracranial or non-invasive) and contrasts (e.g., remembered vs. forgotten items, or, successful retrieval of an association) used in a study (Herweg et al., 2020). Indeed, this hypothesis has recently received empirical support in a study investigating the association between memory encoding and theta activity across iEEG and magnetoencephalography (MEG) (Joensen et al., 2023). This suggests that theta power increases are consistently associated with enhanced memory when analysis is restricted to associational contrasts that only include trials on which recall was successful.

In addition to theta oscillations, there are several other neural frequencies that have been mechanistically linked to memory processing in the medial temporal lobes. Prime among these are gamma oscillations (> 30 Hz) which have also been linked to successful memory encoding and retrieval. For example, Fell et al. (2001) were able to show that successful memory formation was associated with increased gamma synchrony between the hippocampus and perirhinal cortex as measured via iEEG. However, the true significance of gamma oscillations in memory processing may lie in their interaction with theta rhythms. For example, theta-gamma phase-amplitude coupling has been associated with successful memory formation with greater retention when gamma oscillations are coupled to the trough of the theta cycle (Lega et al., 2014). Crucially, evidence suggests that long-term potentiation preferentially occurs at specific phases of the theta cycle, indicating a cellular mechanism by which theta-gamma coupling enhances memory (Fell & Axmacher, 2011). Moreover, this coupling seems to be especially important for item-context binding, a defining feature of episodic memory and a process believed to be a core function of the hippocampus. Staudigl & Hanslmayr (2013) demonstrated that when context is manipulated between encoding and test (match vs. mismatch context), both theta power and theta-gamma coupling are associated with item-context binding. However, only theta-gamma coupling was associated with overall item memory suggesting a crucial role for gamma oscillations when representing items and binding them to their context. This finding is in line with a model of theta-gamma coupling put forward by Jensen & Lisman (2005) in which theta cycles represent an overarching context whereas the coupling of gamma bursts to this cycle represents the sequencing of items within that context.

Finally, the desynchronisation of alpha (8 – 12 Hz) and beta (13 – 30 Hz) oscillations has also been associated with successful memory encoding and retrieval (Hanslmayr et al., 2012). In line with the Complementary Learning Systems framework, Hanslmayr et al. (2016) suggested that alpha-beta desynchronisation and theta-gamma synchronisation may reflect differing roles for the neocortex and hippocampus in mnemonic processing. Under this model, the neocortex represents the content of information, and the hippocampus binds this information together. Sensory input leads to a reduction in alpha-beta power by increasing the firing rates of stimulus-specific neurons in the neocortex. Principles from information theory (e.g., Shannon entropy) can be applied to the phenomenon of alpha-beta desynchronisation to infer that this desynchrony may represent the richness of information being encoded in a memory trace (as a dyssynchronous signal can inherently carry more information). Alpha-beta desynchronisation in turn drives stimulus-specific units in the hippocampus to move their firing forward in theta phase so that it occurs earlier in the gamma cycle. This leads to preferential long-term potentiation for these stimulus-specific hippocampal units resulting in the formation of a hippocampal memory trace for the information currently represented in neocortex. Griffiths et al. (2019) recently provided evidence for this model by demonstrating that decreases in neocortical alpha-beta power precede hippocampal gamma power increases by ~150 ms during memory encoding. Interestingly, it also appears that this relationship is bidirectional, with hippocampal gamma preceding neocortical alpha-beta decreases by ~250 ms during memory retrieval.

In sum, episodic memory is critically reliant on the medial temporal lobes and is underpinned by neural oscillations in the theta, gamma, alpha and beta bandwidths. Importantly, the medial temporal lobes do not function independently, episodic memory is highly reliant on medial temporal interactions with the neocortex. However, I am yet to discuss how schemas fit into this equation. Prior research suggests that areas of the prefrontal cortex play a crucial role in representing schemas and facilitating their influence on memory processing. As such, in the next section I discuss existing models on medial temporal-prefrontal interactions and how these relate to episodic memory and schemas.

1.4 Neurocognitive models of PFC-MTL interaction in memory

As has been discussed, the medial temporal lobes and hippocampus play a key role in the formation and consolidation of memories, especially through interactions with the neocortex. Perhaps the most prominent model of hippocampal-neocortical interaction is Systems Consolidation Theory which proposes that memory formation occurs through a process of rapid learning in the hippocampus followed by longer-term consolidation into the neocortex. Via this process, more remote memories become less dependent on the hippocampus with time, and develop into interconnected neocortical representations independent of the original hippocampal trace.

However, since the conception of this model, research has revealed many findings that do not neatly adhere to the process this model describes. For example, there are numerous accounts of the acquisition of new information in patients with anterograde amnesia. Patient K.C., for example, was found to be able to be show priming effects, new semantic learning, and a rudimentary form of episodic learning referred to as mnemonic precedence in which he was able to remember the location and identity of an object hidden by an experimenter (Rosenbaum et al., 2005; Schacter et al., 1986). Similar findings have also been observed in H.M., who was able to recall the names of, and semantic information about, individuals who had acquired fame after the onset of his amnesia. For instance, H.M. was able to identify Lee Harvey Oswald as the individual who "assassinated the president" (O'Kane et al., 2004). Although it may be argued that these cases of new learning in amnesics primarily concern forms of implicit and semantic learning, rather than episodic memory, it must be noted that recent reviews have provided a strong challenge to the notion that the hippocampus' function in memory is strictly restricted to episodic learning alone (Duff et al., 2020). In a different vein, the standard model of systems consolidation has also been criticised for offering no explanation for how the process of semanticisation occurs nor of how the retrieval of neocortical representations becomes hippocampus-independent despite the hippocampus' initial necessity (Nadel & Moscovitch, 1997). Findings and criticisms of this sort have led to the suggestion that the standard model of systems consolidation may need to be revised.

One way of accounting for some of these discrepancies involves the role that prior knowledge, such as schemas, may play in augmenting learning. A key region of the neocortex implicated in this process is the medial prefrontal cortex (mPFC). The importance of the mPFC in mediating the influence of prior knowledge on new memories can be seen in studies of mPFC lesions. Early work in this area revealed that ventral mPFC (vmPFC) damage could lead to confabulation and proposed that this was a result of deficits in the strategic retrieval of context appropriate memories (Moscovitch, 1997). That is, the mPFC may play an important role in monitoring the relationship between current inputs and existing knowledge. Future lesion studies corroborated this idea by demonstrating that schema reinstatement and instantiation are impaired in mPFC patients (Ghosh et al., 2014; Giuliano et al., 2021). Interestingly, damage to the mPFC can also reduce false memory for new items that resemble past experiences, further indicating a role of the mPFC in monitoring the congruence of inputs with prior knowledge (Spalding et al., 2015; Warren et al., 2014). Taken together, these findings suggest that how new information is integrated into memory may critically depend on its relationship to existing schemas.

The SLIMM framework (van Kesteren et al., 2012), for example, predicts that new learning can be fast tracked into the neocortex if an input is schema-congruent. Indeed, this idea is consistent with the notion of why humans may possess two complementary learning systems in the first place, to prevent the catastrophic interference that can occur when novel (and potentially contradictory) information must be integrated with well-established knowledge. If new information is highly schema-congruent, the drawn out process of hippocampal consolidation that protects against interference may simply not be necessary (McClelland, 2013).

SLIMM proposes that there is a division of labour between the medial prefrontal cortex (mPFC) and medial temporal lobes (MTL) during memory formation, depending on whether incoming information is congruent or incongruent with existing schemas. Under SLIMM, the mPFC acts as a schema congruency detector that “resonates” with information that is congruent with existing knowledge. This resonance results in the inhibition of hippocampal encoding and allows schema-congruent information to be immediately integrated with existing neocortical knowledge structures to which it already relates. Importantly, this fast tracking is predicted to result in a gist-like representation which lacks detailed information regarding the incidental details of the encoding context. On the other hand, when incoming information is incongruent with existing knowledge, SLIMM predicts this will result in detailed hippocampal encoding including incidental features of the encoding context. As already discussed, SLIMM predicts a U-shaped function of schema congruency. This occurs because information that is unrelated, or weakly related, to existing schemas does not trigger strong resonance in the mPFC nor a strong prediction error to drive hippocampal encoding.

SLIMM has received empirical support particularly in terms of its prediction regarding detailed vs. gist-like encoding depending on the schema congruence of an input. For example, Frank et al. (2020) demonstrated that foils that were perceptually similar to to-be-remembered targets were correctly rejected with more accuracy when the target had appeared in an unexpected context during encoding, indicating that the target had been encoded with more detail in the unexpected context. In a similar vein, a different study found that mPFC lesions could prevent false alarms to foils that were perceptually similar to targets encoded under schema-congruent conditions (Spalding et al., 2015). This suggests that the lesions in these patients prevented the mPFC from inhibiting detail-rich encoding for schema-congruent events.

A defining feature of SLIMM is the notion that the mPFC and MTL play distinct, competitive roles in memory formation depending on how incoming information relates to existing knowledge. In contrast to this, there are several models of mPFC-MTL interaction that suggest that the relationship between these two regions is much more complementary. One such framework has been drawn from Trace Transformation Theory. Trace Transformation Theory predicts that the hippocampus is principally involved in the representation of precise perceptual details of a memory and that remote episodic memories can remain hippocampally dependent provided the memory is episodically rich. Robin & Moscovitch (2017) proposed that the ventromedial prefrontal cortex (vmPFC) and subdivisions of the longitudinal axis of the hippocampus represent different features of an experience with schemas being represented in the vmPFC, and gist and details being represented in the anterior and posterior hippocampus respectively. Under this model, the nature of vmPFC-HPC interactions is determined by how an input relates to the information each region represents and to the goals of the memory task. That is, which regions of this network are recruited, and how they interact, depends on the detail-richness of the cues and tasks employed.

Another formulation of mPFC-MTL interaction comes from Preston & Eichenbaum (2013) who suggest that there is a bidirectional relationship between the hippocampus and the mPFC. In this model, there is likewise subdivisions along the longitudinal axis of the hippocampus with the posterior hippocampus (pHPC) coding for objects and the location of the contexts they occurred in, and the anterior hippocampus (aHPC) differentiating between different contexts and linking intra-context events. Context defining information is sent from the hippocampus to the mPFC where distinct contextual rules are developed and represented. When a context is re-encountered, context information is sent from the aHPC to the mPFC which then feeds back appropriate contextual rules to the hippocampus facilitating memory encoding and context-dependent memory retrieval.

A final model that proposes a more complementary relationship between the hippocampus and mPFC was proposed by Takehara-Nishiuchi (2020). A crucial feature of this model is that it extends the use of existing knowledge to facilitate memory encoding from occurring across multiple events to also occurring within a single instance. The hippocampus encodes continuous changes in the spatial and temporal context of an ongoing event. Meanwhile, the mPFC monitors whether a given input is related to the input perceived a moment earlier and integrates these related inputs. Thus, traces are generated within the mPFC that represent the central content of an event. These traces can then exert top-down control over which elements of the detail-rich hippocampal representations are perceived as most salient and this saliency signal can act to potentiate memory encoding of those features. Importantly, this model makes some predictions regarding mPFC-MTL interaction which directly contradict predictions made by SLIMM. Notably, whereas SLIMM proposes that the mPFC resonates with schema-congruent information, the Takehara-Nishiuchi (2020) model states that the mPFC is continuously tracking and integrating commonalities across inputs over time and that this process occurs regardless of schema congruency. Moreover, in stark contrast to SLIMM, Takehara-Nishiuchi suggests that the mPFC plays a pivotal role in potentiating, rather than inhibiting, HPC encoding of behaviourally relevant content.

In sum, precisely how mPFC-MTL interactions govern the schema-mediated formation of memories is still a topic of debate with several relevant models existing in the literature. A key distinction between these models is whether they propose a distinct and competitive relationship between the hippocampus and prefrontal cortex or whether this relationship is more complementary in nature. Nevertheless, there is a consensus that interactions between the regions play an important role in mediating memory formation when schematic knowledge is engaged.

Thus far, we have seen that the representation and monitoring of schemas, and the influence they exert on memory retention, seems to be governed by interactions between the medial temporal lobes and medial prefrontal cortex. However, it is important to consider what other processes may contribute to, or indeed govern, these PFC-MTL interactions. As has already been mentioned, one advantage of possessing schemas is that they can be used to make predictions concerning future events. Indeed, the notion of the hippocampus as a violation detection device is central to several models of mPFC-MTL interaction. As such, these interactions may be modulated by a predictive coding process by which the brain maintains a generative model of the environment. Next, I discuss predictive processing theories in more detail and assess how they may be linked to the emergence of schema effects in memory.

1.5 Predictive processing and memory

Predictive processing theories suggest that a core function of the brain is to create a generative model of the environment by which predictions regarding expected sensory inputs can be generated with the aim of minimising prediction errors across time. Sensory information and prior experience are used to generate predictions about what is likely to occur in a given context, which are then compared with incoming sensory perceptions. The generation and updating of this model based on prediction errors consequently allows an organism to perceive, learn and adapt its behaviour in an ever-evolving environment. First formally described for use in the neurosciences in the late 1990's and mid 2000's (Friston, 2005; Rao & Ballard, 1999), predictive coding has become highly influential in the field as it offers a potentially unifying account of cortical function that can model brain function at all three of the levels of analysis set forward by Marr (1982) (computational, algorithmic and implementation). Mathematically, predictive coding theories are Bayesian frameworks as they describe the updating of an internal model of the environment based on a comparison of new information with prior knowledge (Friston, 2005; Millidge et al., 2021). Under such a framework, a prior distribution represents an organism's predictions based on existing knowledge and context, and an evidence-based likelihood distribution represents incoming sensory information. Here, prediction error is calculated based on the divergence of the prior and likelihood distributions resulting in a posterior distribution that represents updated knowledge.

Predictive coding theories were originally developed to model perceptual processing but have also been successfully applied to the field of learning and memory. One such model of memory is the Predictive Interactive Multiple Memory Systems (PIMMS) model (R. N. Henson & Gagnepain, 2010) which proposes that perceptual predictive coding can account for phenomena in recognition memory. PIMMS builds upon previous theories of multiple memory systems by asserting the existence of three fundamental memory systems: episodic, semantic, and perceptual systems that are primarily situated in the hippocampus, perirhinal cortex and posterior sensory cortices respectively. Perceptual information feeds forward from posterior sensory cortices, via the ventral visual stream, to the perirhinal cortex and subsequently the hippocampus.

PIMMS distinguishes itself from other mnemonic models by proposing that, rather than processing occurring separately in each system, encoding and retrieval are a consequence of recurrent interactions between the systems with both feedforward and feedback connections. The role of feedback mechanisms is to predict the patterns of activity in the preceding level of the system. For example, the hippocampus may use contextual representations to predict what items are likely to occur in that context, with those items being represented in the preceding semantic system in the perirhinal cortex. These predictions can then be compared with incoming feedforward connections to determine their accuracy. Such a recurrent flow of information can thus be used to minimise prediction errors over time, facilitating perception and recognition. Notably, PIMMS primarily concerns itself with the processes of recollection and familiarity in recognition memory and how predictions in one memory system can help to facilitate encoding and retrieval in another. Importantly, PIMMS does not make any explicit predictions concerning how temporal predictions of future events are implemented in the brain and can be used to facilitate episodic encoding or retrieval more generally. Nonetheless, there is empirical evidence in support of PIMMS (Greve et al., 2017) and it provides a good example of how predictive processes may play an important role in human memory.

A key finding in research on the hippocampus is that it appears to be highly sensitive to prediction errors and may be the hub of our internal generative model (Barron et al., 2020; J. Lisman & Redish, 2009; Sinclair et al., 2021; Stachenfeld et al., 2017). This generative model could allow for temporal predictions of future events that in turn facilitate the encoding of behaviourally relevant information. Indeed, fMRI studies have demonstrated that hippocampal activity is increased after prediction violations with the CA1 region seeming to be of particular importance (Bein et al., 2020; Chen et al., 2015). There is a clear link here with certain models of mPFC-MTL interaction which predict that the hippocampus is specifically involved in the acquisition of schema-incongruent knowledge (van Kesteren et al., 2012). When one links these predictive processes with models of mPFC-MTL interaction a potential model of schema effects emerges. As we experience an event the brain is continuously predicting what might happen next. The affirmation or violation of these predictions is then associated with activity changes in the mPFC-MTL axis, which subsequently result in the behavioural manifestation of schema effects in memory. Although some evidence for this model exists (Sinclair et al., 2021), evidence for each stage of this process, from ongoing neural predictions to behavioural schema effects, is under explored. Moreover, the frequent use of fMRI in previous studies leaves questions regarding the temporal dynamics of these processes underspecified.

SLIMM is the main model that explicitly states how schema effects could emerge from prediction errors triggering changes in mPFC-MTL interaction. However, some evidence from the prediction violation literature may be at odds with SLIMM's predictions here. For example, Garrido et al. (2015) combined a sequence learning paradigm with MEG and dynamic causal modelling (DCM). Sequences could fall into predictable, mismatch (prediction error), or unpredictable conditions. Source reconstruction suggested that prediction error signals were associated with theta-band activity in the hippocampus and vmPFC. However, DCM suggested that hippocampal theta oscillations were driven by vmPFC theta. This is at odds with SLIMM which suggests that the vmPFC should be relatively inactive under schema-incongruent conditions. Importantly, however, Garrido et al. (2015) did not take measures of subsequent memory, and therefore these vmPFC-HPC interactions cannot be assumed to have triggered the generation of behavioural schema effects. Nonetheless, such research highlights the potential importance of effective connectivity measures in distinguishing between models of mPFC-MTL interaction and how they are related to predictive coding processes. Crucially, there is a body of evidence to suggest that the hippocampus plays an important role in prediction error detection and that this may be linked to subsequent patterns of activity in the mPFC-MTL axis that predict successful memory encoding.

Recent studies have attempted to address whether predictive cues can trigger the reactivation of representations before the predicted items actually occur. One such study comes from Kok et al. (2017) who combined magnetoencephalography with multivariate decoding. Participants heard predictive audio cues that predicted the orientation of an upcoming grating. Classifiers trained to decode these orientations were then applied to the pre-stimulus period between the audio cue and the onset of the grating. Orientation could be decoded up to 40 ms before stimuli onset when the audio cue accurately predicted the upcoming orientation. This indicates that prior expectations can indeed engage the pre-activation of sensory templates. Interestingly, a follow up study (decoding shapes instead of gratings) found evidence of pre-stimulus expectation not in overall classifier performance, but in the frequency at which decoding oscillated across time (Hetenyi et al., 2024). A Fast Fourier Transform was applied to the pre-stimulus decoding time series to assess its power spectrum. Relative to a classifier that was trained on randomised labels, it was found that the shape classifiers had greater power in the 10 - 11 Hz alpha power range. Importantly, this suggests the pre-activation of upcoming items can be detected in the brain, and this may manifest both in overall classifier performance and in the spectral power at which a decoder oscillates. However, once again, these studies did not assess how these predictions, and their affirmation or violation, might be related to subsequent memory for the predicted items.

Finally, another concept related to neural predictions that has relevance to the current thesis is the concept of predictive value. Predictive value refers to the propensity of a stimulus to predict future events. If A predicts B, then A is said to have predictive value. Crucially, evidence suggests that when an item has high predictive value it may impede its own encoding into memory. Sherman & Turk-Browne (2020) used a sequence learning paradigm in which participants viewed a continuous sequence of scene images and made judgements as to whether the image contained a man-made object or not. The scene images came from 12 different categories of scenes and were split into three conditions: a predictive condition, in which the image reliably predicted the following category, a predicted condition, in which the category was reliably predicted by the preceding image, and a control condition, in which the scene was neither predictive nor predicted. After an encoding phase there was a surprise old-new recognition test. It was found that the hit rate for the scenes in the predictive condition was significantly lower than the hit rate for the control images. Sherman and Turk-Browne hypothesised that this occurred due to the co-dependency of statistical learning and episodic memory on the hippocampus. More specifically, the predictive image triggers the retrieval of a representation of the predicted category which in turn inhibits the encoding of the details of the original predictive image. Support was lent to this hypothesis in an fMRI experiment and a follow up iEEG study

(Sherman et al., 2022) with fMRI data indicating that the extent of impaired memory for predictive items is associated with the reinstatement of information regarding the upcoming category in the hippocampus. This finding may have implications for how prediction errors can facilitate encoding in real-world situations. Schema effects suggest that memory is enhanced for schema-congruent well-predicted events. However, in the real-world it is often the case that we experience a chain of predictive events with a formerly predicted event also predicting the next in a sequence. It is unclear whether schema congruency effects could emerge in a situation in which all sub-events are both predictable and predictive, an occurrence which is common in everyday experience.

In sum, predictive coding may contribute to the emergence of schema effects in memory by modulating mPFC-MTL interactions based on whether ongoing predictions are affirmed or violated. However, to date, no studies have addressed whether ongoing neural predictions can be linked to the emergence of behavioural schema effects. Addressing this gap in the literature will be a primary aim of the current project. However, before formally laying out my research questions, I will discuss the neural basis of sequence learning and temporal order memory, as understanding these processes is key to the investigation of how temporal predictive coding processes may contribute to schema effects in memory.

1.6 Temporal order memory and sequence learning

Real-world experiences are composed of sequences of events that unfold over time. As a result, a core component of episodic memory is the representation of temporal order. Importantly, the preservation of temporal order is crucial if one is to use past experiences to make predictions of the future. Thus, temporal order information is likely a core component of many schemas. Indeed, previous research has already demonstrated that prior knowledge may be used to facilitate the perception of the scenes we encounter across time. M. E. Smith & Loschky (2019) presented participants with sequences of scenes that either formed a coherent spatiotemporal route or were presented randomly. Categorisation accuracy was greater for scenes that were presented in spatiotemporally coherent sequences and when the preceding scene shared visual characteristics with the target scene. Moreover, a second experiment demonstrated that participants could more accurately predict upcoming scenes in coherent sequences, and that both predictability and the visual similarity of sequential scenes independently predicted categorisation performance. Together, these experiments suggest that predictions based on prior knowledge facilitate the accurate recognition of scene-gist. Moreover, a subsequent event-related potential (ERP) study found reduced amplitudes for both the N300 and N400 components for coherent sequences (M. Smith et al., 2021). These components are associated with the recognition and categorisation of visual stimuli (N300) and with semantic integration based on expectedness (N400). Thus, their reduced amplitude further corroborates the conclusion that ongoing predictions during coherent scene perception supported efficient perceptual processing. Importantly, however, although these findings are in line with a predictive coding account of perception that uses schemas to facilitate temporal predictions of upcoming events, this research did not address how these predictions influenced subsequent memory.

Many theories of how the mind retains serial order have their origins in the short-term memory (STM) literature. One obvious model of sequence order retention is to assume that each element in a sequence is associated with the next in a pairwise manner, a process known as chaining. However, chaining has been largely refuted as a potential mechanism for sequence learning in STM models as it is not consistent with observed patterns of error in immediate serial recall (R. N. A. Henson et al., 1996). Instead, certain models of STM serial order recall involve associating each item in a sequence with a “context” signal that evolves over time, generating different states that can be associated with the different items in a sequence (Burgess & Hitch, 2005). Alternatively, other models code the items of a sequence relative to their position with regards to the start and/or end of the sequence (R. N. A. Henson, 1998; Page & Norris, 1998). However, the relevance of these processes to understanding serial order retention in long-term memory is still a topic of debate. Research into the Hebb effect, by which repeated exposure to a sequence enhances future serial recall of that sequence, suggests that a drifting context-based model may indeed have relevance to understanding long-term sequence learning (Hitch et al., 2009). Furthermore, and in contrast to STM models, single-cell recordings in rats suggests that chaining may be applicable to the learning of sequences in long-term memory (J. E. Lisman et al., 2005).

Importantly, the hippocampus has been demonstrated to be an important neural structure for sequence learning and the representation of temporal order. For example, MTL activity has been shown to increase during successful encoding of associations presented across a time delay (Hales & Brewer, 2010) and rats with hippocampal lesions seem to present difficulties with remembering the temporal ordering of associations (Fortin et al., 2002; Honey et al., 1998). Interestingly, some research suggests that the hippocampus may use time cells to generate a gradually evolving context representation, much like the one described to explain serial order recall in STM, against which each item in a sequence can be associated (Howard & Eichenbaum, 2013; Howard & Kahana, 2002).

Davachi & DuBrow (2015) have highlighted a key distinction between two different forms of sequence learning and how they may be implemented in the hippocampus. These two forms of sequence learning are referred to as *sequence learning through repetition* and *episodic sequence encoding*. As the name suggests, sequence learning through repetition refers to the learning of a temporal relationship that occurs due to repeated exposure to a given sequence structure, with or without explicit awareness. Evidence suggests that repeated exposure in this way may lead to the generation of a predictive code in the hippocampus. When a sequence is encountered numerous times, enhanced hippocampal activity can be observed during sequence perception. This increase in activity has been hypothesised to be indicative of hippocampal pattern completion, the process by which a partial cue from an episode can result in the holistic recollection of the entire event (Horner et al., 2015; Marr, 1971; Nakazawa et al., 2002; Rolls, 2013). Thus, prior exposure to the sequence facilitates the prediction of the upcoming items via the hippocampus. Evidence for this has been provided by combining sequence learning paradigms with fMRI pattern similarity analyses. Hsieh et al. (2014) presented participants with sequences of visual objects that could occur either in a previously presented order or in a random sequence. Hippocampal representational similarity was shown to be increased when objects were presented in their learned sequential positions relative to the same objects presented in random positions or different objects in the same position. This finding indicates that sequence learning is associated with representational change of the objects in a sequence depending on their sequential context. Importantly, another study indicates that this representational change occurs in a feedforward manner throughout a sequence. Schapiro et al. (2012) presented participants with a continuous stream of fractal images in which given fractals predicted the occurrence of one another. It was found that the direction of pattern similarity in the CA2, CA3 and dentate gyrus subregions of the hippocampal formation was asymmetrical between strongly paired items. That is, the representational pattern of the first item of a pair became more similar to the pattern of the second item than did the second item's pattern to the first. Thus, hippocampal pattern similarity changes during the learning of repeatedly presented sequences seems to be representative of a forward prediction. In sum, this evidence suggests that the prediction of the upcoming items in a sequence via hippocampal pattern completion may play an important role in sequence learning through repetition.

Episodic sequence encoding contrasts with sequence learning through repetition in that it refers to learning of a sequence after a single exposure. It has been hypothesised that episodic sequence encoding may involve the use of a temporal context representation as described above. For example, the Temporal Context Model (TCM) (Howard & Kahana, 2002) proposes that during episodic sequence encoding each element in a sequence is associated with a different state of a drifting context representation. During recall, the retrieval of a given item in a sequence can result in the reinstatement of its neighbouring items as they share a similar context state. Importantly, TCM predicts that the context state of item N in a sequence is encoded in the context state of N + 1 but not vice versa. This allows for the representation of the sequential ordering of events and can help to explain contiguity effects by which items following a cue are more likely to be recalled than items that preceded it (Kahana, 1996). Thus, episodic sequence encoding may be biased toward a different mechanism for representing temporal order than would sequence learning through repetition. However, it is noteworthy that these two processes are by no means mutually exclusive. Indeed, a drifting context representation may play an important role in sequence learning through repetition in as much as shared contextual features are also repeated. Likewise, hippocampal pattern completion of subsequent items in a sequence may begin to occur after only a single exposure depending on how strongly the relationships between items were encoded.

Finally, it is worth noting that theta rhythms also play an important role in the representation and retrieval of temporal order information at a neural level. For example, EEG studies indicate that theta power is increased during the active maintenance of temporal order information during working memory tasks (Hsieh et al., 2011). One interpretation of this data is that theta-band activity represents online temporal order information. However, another possibility is that this theta activity represents prediction errors during sequence learning which would reduce as sequence structure becomes more familiar. Crivelli-Decker et al. (2018) used EEG to try and tease apart these two possible interpretations. Participants were exposed to repeated sequences of visual items that either appeared in a consistent or a random order across repetitions. It was found that theta power decreased across sequence learning for items in the consistent sequences and the extent of these power decreases correlated with subsequent reaction times for semantic judgements on upcoming items. These findings are consistent with the notion that theta power is associated with the use of prior knowledge to try and predict upcoming events in a sequence with greater power representing more prediction error. This finding also provides support for the Davachi & DuBrow (2015) theory surrounding sequence learning through repetition in which pattern completion facilitates the prediction of upcoming events and suggests that successful prediction in such cases is associated with decreases in theta power.

Another role that theta appears to play in sequence learning relates to phase precession, in which the sequencing of locations in space are phase-locked to the theta cycle. Importantly, a recent study extended this phenomenon from rodents to humans and demonstrated that it generalises to the conceptual domain. Reddy et al. (2021) presented participants with sequences of visual images while recording iEEG data from the hippocampus and the temporal cortex. It was found that neuronal spikes for consecutive images in a sequence were phase-locked to progressively earlier phases of the underlying theta cycle. Interestingly, this suggests that the neural process used to represent sequencing in physical space is the same as that used to represent sequencing in conceptual space. However, given the fine spatial resolution required to index the neurons or populations of neurons which are representing individual items in a sequence, phase precession of this kind cannot be easily indexed using neuroimaging methods such as MEG. In sum, theta oscillations seem to be associated with sequence learning in two ways: Theta power appears to represent the expectedness of items in a sequence across learning and theta phase precession seems to code for temporal order information.

Before finishing, it is worth highlighting that sequence learning features prominently in the implicit statistical learning literature (Destrebecqz & Cleeremans, 2001; Frost et al., 2019; Schapiro & Turk-Browne, 2015). Although interesting, a detailed focus on the distinction between explicit and implicit processes in sequence learning is beyond the scope of the present thesis and will only be addressed when discussion of this issue is pertinent to our primary hypotheses regarding schema, temporal order and memory encoding.

1.7 Thesis Overview

The preceding review highlights a potential model of how schemas influence memory processing. As we experience the world, we develop schemas that represent statistical regularities across multiple episodes. These schemas can then form a part of our internal generative model, allowing us to make predictions about upcoming events. The affirmation or violation of these predictions can then be used to indicate what information should be encoded into memory, so as to update our internal model as the environment evolves. Evidence suggests that interactions in the mPFC-MTL axis play an important role in such memory encoding.

Crucial to this theory of mnemonic function is the notion that we use schemas to make predictions about upcoming events. However, to date, most studies of schema effects in memory have used paradigms that do not require temporal prediction of the future. As such, the present thesis aims to see if schema effects can be observed when a key feature of a schema is the chronology of events. Another key feature of the above theory is that schema effects should be preceded by ongoing neural predictions and, where appropriate, followed by prediction error signals. To Investigate this, I will combine a temporal order memory paradigm, in which participants acquire schematic knowledge of sequence structure, with magnetoencephalography (MEG) and multivariate pattern analysis (MVPA) in an attempt to see if predictive reactivation of schematic knowledge and subsequent prediction errors can be linked to the behavioural manifestation of schema effects in episodic memory.

Research questions:

1. Can behavioural schema effects (congruent and incongruent) be indexed in a temporal order memory paradigm in which a core feature of the schema is the chronology of events?
2. Can MEG and MVPA be combined to index predictive reactivation of upcoming items in a sequence after participants have been trained on the sequence's schematic structure?
3. When predictions are violated, is there evidence for behavioural and neural prediction error signals associated with enhanced memory retention of schema-incongruent information?

2 Schema Effects in Temporal Order Memory

2.1 Introduction

Schemas are associative knowledge structures that represent the semantic knowledge we acquire through repeated exposure to events that share a common structure. By extracting statistical regularities across multiple episodes, this schematic knowledge can be used to facilitate adaptive behaviour by generalising from past experiences to present events. In addition to the obvious advantage of being able to generalise across events, it is also well established that schemas can augment new learning. That is, memory retention appears to be enhanced for events that are congruent with an individual's existing schemas (Bein et al., 2015; Buuren et al., 2014; Mandler & Johnson, 1977; Raykov et al., 2019; Staresina et al., 2008; Tse et al., 2007; van Kesteren et al., 2010). Moreover, and somewhat paradoxically, a memory advantage also seems to be present when events are schema incongruent (Greve et al., 2017; Kafkas & Montaldi, 2018; Mäntylä & Bäckman, 1992; Tulving & Kroll, 1995).

Despite some research showing that schema-congruent information boosts memory, whereas some studies show that incongruency boosts memory, recent research has established that it is possible to index both a congruency and an incongruency benefit in a single paradigm. Greve et al. (2019) proposed that a core issue in schema research is the frequent conceptualisation of congruency as a binary, rather than a continuous variable. Real events cannot be neatly categorised into congruent vs. incongruent instances, and are more accurately conceptualised as falling on a spectrum of schema congruency. Under this model, in addition to an event being either congruent or incongruent with a schema, it is also possible for an event to be unrelated to a schema, in that it does not clearly relate to existing schemas but neither does it actively contradict that knowledge. In such an unrelated case, memory may be weakened as compared to events that are schema-congruent or incongruent. As such, memory retention may be modelled as a U-shaped function of schema congruency, with poorly encoded schema-unrelated events found in the centre of the congruency spectrum.

Greve et al. (2019) provided evidence for this model in a recognition memory paradigm. During study, participants were presented with differing numbers of two types of objects (e.g., one apple, two teddy bears). Across four trials per pair of objects, participants were required to learn, via trial-and-error plus feedback, which object had the highest value after being told that one object had twice the value of the other. Crucially, across the four trials the rule was either predictable (congruent condition; the rule was the same across all four trials), unpredictable (unrelated condition; no consistent rule was ever established), or the rule was flipped midway through the trials (incongruent condition; the rule was initially predictable but was then changed). Recognition memory was then assessed for the displays of objects viewed during study. It was found that recognition memory was enhanced for displays in the congruent and incongruent conditions relative to the unrelated baseline in which no schema was established. This study was the first to demonstrate both a congruency and incongruency advantage in a single paradigm. However, schemas in this study (the value rule) were generated in a single session, shortly before testing. Schemas are generally thought to be well-consolidated neocortical representations acquired across multiple episodes, and it is questionable whether the small number of trials in this study, with no intervening consolidation period, would allow for the generation of such a representation.

Quent et al. (2022) were able address this issue by demonstrating a similar U-shaped curve using spatial schemas that pre-existed due to frequent real-world experiences. Participants explored a virtual reality kitchen environment and memorised the locations of 20 objects. They then re-entered the emptied kitchen and placed the objects in their previously observed locations. It was found that object-location memory was greatest for objects that were subsequently rated to have been observed in locations that were most expected or unexpected.

Together, these studies provide good evidence that the use of a schema-unrelated baseline condition may be an important step when trying to distinguish between the effect of schema congruency and incongruency on subsequent memory. In the present chapter, I ask whether a similar U-shaped function can be observed for other forms of memory (e.g., temporal order memory) and when schemas encode for additional information other than contextual rules or spatial organisations (e.g., chronology). Next, I will discuss why this is an important issue to consider.

Episodic memories are defined by their event-based nature and the spatiotemporal structure of their content (Tulving, 2002). Not only do episodic memories encode when and where an event occurred but also the temporal ordering of events. As such, Ghosh & Gilboa (2014) have highlighted that some schemas should have sensitivity to chronology with the temporal relationships between units being encoded in a schema's associative network. Furthermore, it may be that even within a single schema, multiple temporal orders may need to be represented. For example, a schema for a restaurant may consist of nodes connecting related concepts such as getting a table, reading a menu, ordering and receiving food, and paying the bill. However, the temporal relationship between these nodes may differ based on context. When visiting a traditional restaurant, for example, it is typical to wait at the door to be seated, order your food from the table, to receive your food, and then to pay the bill. However, in a different setting such as a pub, it is more typical to take a seat at any table that is free, order your food at the bar, pay immediately, and then to sit back down to receive your food. Thus, we see that the temporal ordering of events is a defining feature of episodic memories and, as such, chronological information should be a core feature of many schemas. Indeed, a major adaptive advantage of possessing schemas in the first place is that they may facilitate temporal predictions of the future, allowing an individual to anticipate how an event may unfold. In fact, schema effects in general may be a consequence of predictive coding processes (Friston, 2005; R. N. Henson & Gagnepain, 2010). Under such a model, prior knowledge is used to make predictions (e.g., representations of upcoming stimuli may be pre-activated prior to their occurrence) and memory encoding and/or retrieval is modulated based on the affirmation or violation of these predictions.

A neurocognitive model that implicates predictive processing in the emergence of schema effects is the schema-linked interactions between medial prefrontal and medial temporal lobe framework (SLIMM) (van Kesteren et al., 2012). Under SLIMM, there is a division of labour between the medial prefrontal cortex (mPFC) and medial temporal lobes (MTL) during memory encoding dependent on whether an input is congruent with our existing knowledge. The mPFC is said to "resonate" with inputs that are congruent with existing representations in the neocortex. If this occurs, the mPFC inhibits hippocampal encoding so that this information can be immediately integrated into related neocortical stores. However, if an input is incongruent with our existing knowledge this triggers a prediction error in the MTL that potentiates hippocampal encoding. Critically, when an input neither resonates with the mPFC nor is surprising enough to trigger a prediction error (i.e., it is unrelated to existing knowledge), neither the mPFC nor the MTL is particularly activated leading to poor memory encoding.

SLIMM also predicts differences in the detail-richness of newly encoded representations dependent on how they relate to existing knowledge. Schema-congruent memories that were encoded via the mPFC are predicted to be gist-like, whereas schema-incongruent memories encoded by the MTL are believed to be more detailed and include incidental details regarding the encoding context. Notably, although SLIMM states that prediction errors can drive learning, it does not state that representations of upcoming events/items are necessarily pre-activated prior to their actual occurrence. Rather, incoming inputs may simply have high or low resonance with existing neocortical knowledge, and this alone can modulate encoding. Nevertheless, whether predictive coding processes generate schema effects via temporal predictions of the future is an interesting avenue to explore as it would provide valuable information regarding the circumstances under which, and the mechanisms by which, schemas can modulate memory.

Crucially, research on schema effects to date has tended to use schemas that are somewhat static, lacking a temporal element to their structure and instead focusing on contextual rules, associative relationships, or spatial organisations. If schema effects are indeed driven by predictive coding processes and schemas in general are sensitive to chronology, then schema effects should be present for memory of temporal order and when a core feature of a schema is the chronology of events. Indeed, there is already good evidence that predictions are routinely employed during sequence learning and perception. For example, the spiking of hippocampal place cells has been shown to represent locations that are ahead of a rat traversing a maze, and there is evidence to suggest that this process extends beyond the spatial domain and may also represent points in time or the traversing of a sequence of internally generated states (J. Lisman & Redish, 2009). Furthermore, the perception of scene-gist can be facilitated when scenes are perceived in a spatiotemporally coherent order, and research suggests that the predictability of an upcoming scene contributes to this effect (M. E. Smith & Loschky, 2019). Moreover, this effect appears to reduce the amplitude of the N400 event-related potential (M. Smith et al., 2021). The N400 has been linked to semantic integration and has a heightened amplitude for semantically incongruent or unexpected content. Therefore, its reduced amplitude for predictable scenes supports the idea that ongoing predictions of upcoming events can aid efficient perceptual processing. However, although the findings of these studies are in line with a predictive coding account of perception that makes use of prior knowledge to make temporal predictions of the future, this research has not addressed how these ongoing predictions influence subsequent memory or could lead to the generation of mnemonic schema effects.

In the following experiments, I use a temporal order memory paradigm to investigate whether schema effects can emerge when a core feature of a schema is the temporal order of events that could allow for temporal predictions of the future. The studies are split into two sessions separated by 24 hours. During session 1, participants pre-learned *sequence schemas* which related to the order of semantic categories within a given context. During pre-learning participants also made animacy judgements on the categories which introduced a categorisation and motor response element to the schemas. During session 2, participants repeated the pre-learning task before moving on to the main temporal order memory task. The main task was split into study-test blocks. At study, participants were presented with new sequences that used novel exemplars (different from those used at pre-learning) but that either affirmed, violated or were unrelated to the pre-learned schemas. While studying these novel sequences participants continued to make animacy judgements, and response times and accuracy were assessed to determine whether pre-learned knowledge was being used to make predictions of upcoming events. At test, participants' memory for temporal order was assessed in a cued recall four-alternative forced-choice task. Participants were cued with an exemplar from a sequence and were required to retrieve the following two items, that I refer to as Cue+1 and Cue+2 respectively. Notably, I also aimed to distinguish between gist-like and detailed encoding as predicted by SLIMM. This was done by assessing memory at both the category (gist-like) and exemplar level (detailed). In line with Greve et al. (2019), I predicted that temporal order memory should follow a U-shaped function of congruency with memory for schema-congruent and schema-incongruent sequences being enhanced relative to sequences that were unrelated to the schemas.

2.2 Behavioural Experiment 1

2.2.1 Methods

2.2.1.1 Participants

A power analysis was conducted using G*Power 3.1.9.7 (Faul et al., 2007) to determine the required sample size to detect both a schema congruency and schema incongruency effect. Effect sizes were taken from prior literature and assumed $d = .89$ for a schema congruency effect (Greve et al., 2017) and $d = .57$ for a schema incongruency effect (Brod et al., 2015). Power was calculated for a two-tailed paired-samples t-test. Based on these parameters a sample size of $N = 30$ was determined which would give 99% power to detect a congruency effect and 85% power to detect an incongruency effect. Given the novelty of the present paradigm, these effect sizes may not be entirely applicable. However, the data from Experiment 1 was used to calculate more accurate effect sizes for the power analyses of future experiments.

Thirty participants were recruited via SONA systems. Two participants were excluded as they were not significantly above chance at retrieval or encoding as determined by a one-sided binomial test. This left a final sample size of $N = 28$ participants (23 females, 5 males; $M \pm SD$ age in years = 20.16 ± 0.91). All participants were fluent English speakers with normal or corrected-to-normal vision. Participants gave informed consent for the experiment, which was approved by the Research Ethics Committee of the Department of Psychology, University of York. All participants participated for course credit.

2.2.1.2 Materials

The experiment used 196 images of exemplars from each of eight different categories: buildings, dogs, cars, faces, furniture, insects, plants, and tools. Half of these categories can be classified as animate (dogs, faces, insects, plants) and half can be classified as inanimate (buildings, cars, furniture, tools). Face images were taken from the Chicago Face Database (CFD) (Ma et al., 2015). For copyright reasons, face images displayed in figures are AI-generated and are not taken from CFD. A subset of the dog and insect images were taken from the C.A.R.E stimulus set of normed animal images (Russo et al., 2018). All other images were sourced from Google and DuckDuckGo image searches. Images had their backgrounds removed to transparency and were scaled to fit within a 600x600 pixel square.

The eight categories were used to create two six-item sequence schemas for each participant. For each participant, each schema contained the same six categories, three of which were animate and three of which were inanimate. Thus, each participant only ever saw six of the eight possible categories, however, which categories the participants saw was counterbalanced across participants. A schema for each sequence existed at both the category and the animacy level. The schemas at the animacy level for all participants were as follows: (a) *inanimate-animate-inanimate-animate-inanimate-animate*, (b) *inanimate-animate-inanimate-inanimate-animate-animate.* Which categories fell into which positions was counterbalanced across participants, however, a category could never be neighboured by the same categories across the two schemas. For example, if the category *face* was neighboured by *building* and *car* in schema A, *face* could not be neighboured by *building* or *car* in schema B. Finally, each schema was paired with a colour, either red or blue. This colour was used as a context cue at the beginning of each sequence to indicate which schema the participant was about to see. As such, an example of the schemas at the category level for any given participant could be as follows: (red) *tool-dog-building-face-car-plant*, (blue) *building-plant-tool-car-dog-face.*

2.2.1.3 Design

The experiment used a 2x3 (Test x Congruency) within-subject design. The key manipulation was whether a specific sequence was congruent, incongruent, or unrelated to the pre-learned sequence schemas. The independent variable *Congruency* therefore had three levels: Congruent (sequence conforms to a pre-learned schema), Unrelated (sequence is unrelated to any pre-learned schema) and Incongruent (sequence violates a pre-learned schema). The second independent variable, *Test*, had two levels, however, the levels of Test differed between the encoding (study) and retrieval (test) analyses. At encoding, the levels of Test were Cue vs. Cue+1. At retrieval, the levels of Test were Cue+1 vs. Cue+2. The Cue relates to the exemplar prior to any potential schema violation (in the Incongruent condition), Cue+1 relates to the exemplar that potentially violates the schema, and Cue+2 relates to the exemplar following the potential violation.

2.2.1.4 Procedure

A schematic of the experimental procedure can be seen in Figure 2.1. The study was split into two sessions separated by approximately 24 hours. In session 1, participants completed the pre-learning task. Twenty-four hours later in session 2, participants repeated the pre-learning task before moving onto the main temporal order memory task.

2.2.1.4.1 Pre-learning task

During pre-learning participants learned the sequence schemas in an explicit manner, that is, participants were told that the sequence schemas (referred to as “sequence structures” to the participants) existed and that their goal was to try and learn them.

During pre-learning, participants engaged in study-test blocks. During study, participants viewed a single sequence. At the beginning of each trial an empty box appeared in the centre of the screen for 2000 ms. The square was then filled with the colour red or blue for 1500 ms to indicate which schema the participant was about to view. The colour was then removed, and a fixation cross appeared in the centre of the box for 500 ms. Each exemplar was then presented on the screen for 1000 ms each, each separated by a 500 ms fixation cross. While each exemplar was on the screen participants were instructed to make an animacy judgement (animate vs. inanimate) on the image using the “b” and “m” keys which were counterbalanced across participants.

At test, participants were required to recall the order of the categories for the sequence they just saw. To do this, the six exemplars were presented in a row on the top half of the screen in a random order with a coloured rectangle (representing the colour of that schema) below. Participants used the 1-6 keys to reorder the categories into the order that they saw them. Upon reordering the categories, participants received feedback as to whether they had recalled the order correctly or not. If the participant got the order correct, the word “Correct!” appeared in green text for 1000 ms. If the participant got the order incorrect, the word “Wrong!” appeared in red text for 1000 ms.

To complete the pre-learning phase, participants were required to reach criterion and pass a test. At the beginning of the pre-learning phase participants saw three red and three blue sequences randomly intermixed. After this, a counter was started to keep track of how many sequences had been reordered correctly in a row. Reordering a sequence incorrectly reset the counter to zero. Participants were required to get six sequences in a row correct before a test was triggered. Importantly, each six sequences a participant viewed was made up of three blue and three red sequences randomly intermixed. The test was preceded by an instruction screen informing the participants they were about to be tested. The test involved participants reordering both the red and blue sequence as described above but doing so back-to-back and without being shown an example of the sequence first. If they got both test sequences correct, pre-learning was complete. However, if they got one or more of the test sequences incorrect, they were sent back to the criterion task and were required to get another six sequences in a row correct before being tested again. To further reinforce learning upon completion of the test, participants were provided with a pen and paper and were asked to write down the order of the categories for each schema using free recall.

After a 24-hour delay, participants returned to the lab for the second session. The second session began with a repeat of the pre-learning task. At session 2, participants still had to reach criterion and complete a test, however, participants were not required to free recall the schemas using a pen and paper after completion of the pre-learning test.

2.2.1.4.2 Temporal order memory task

After completion of the second pre-learning phase on day two, participants engaged in the main temporal order memory task. The main task was split into 50 study-test blocks. During study the task was identical to that described for the study phase of the pre-learning task, however, rather than seeing just one sequence, participants were presented with three in a row. All sequences in the main task used novel exemplars that had not been seen during the pre-learning task or in any previous sequences in the main task. Crucially, these sequences could fall into any one of the three Congruency conditions. Red and blue sequences could either be Congruent or Incongruent with the pre-learned schemas. Incongruent sequences were initially congruent with the schema, however, on the third or fourth item the order of the remaining categories was altered so that the categories appeared in an unexpected order. In the main task, I also introduced Unrelated sequences that contained the same six categories but presented in a random order. Unrelated sequences were preceded by the context colour yellow. Participants were explicitly instructed that the yellow sequences had no structure to them and that the categories would appear in a random order. Any given block contained one yellow sequence along with either two red, two blue, or a red and a blue sequence. For the red/blue sequences in any block, either both sequences were Congruent, both were Incongruent, or one was Congruent, and the other was Incongruent. In total there were 50 sequences in each Congruency condition. The order in which the three sequences were viewed at encoding was randomised.

At test, temporal order memory was assessed using a four-alternative forced-choice cued-recall task. The test order of the three previously encoded sequences was randomised so that it could differ from the order presented at study. For example, if a participant saw a red, a blue, and then a yellow sequence at study, any of these three sequences could be tested first, second or last at test. On any given test trial participants would see a white square containing a fixation cross in the centre of the screen for 2000 ms. After this, a cue would appear in the box for 1500 ms. The cue was an exemplar from one of the three sequences viewed during the study phase. The cue could either be the second or the third image from the sequence being tested (i.e., the item immediately before the violation in the Incongruent condition). After the cue, four options appeared on the screen (one target, three foils) each contained in a white square presented as a 2x2 grid. Participants were required to select the exemplar they thought came after the cue (Cue+1). Upon selecting an option for Cue+1, another four options would appear, and participants were required to select the image they thought came after the one they had just selected (Cue+2). Participants responded using the “t”, “o”, “b” and “m” keys which represented quadrants on the screen: top-left, top-right, bottom-left, and bottom-right respectively. The locations of the target and foil exemplars on the screen were randomised. For each set of options, participants had 5000 ms to respond before being timed out. A time out resulted in the text “TIME OUT!” being presented in red font for 1000 ms. Time out trials were marked as incorrect. Participants were instructed to prioritise accuracy over speed but, if they were unsure of the correct answer, that it was better to take their best guess than to let the trial time out.

Critically, the foils were set up to allow the distinction of performance at both the category and the exemplar level. If a participant selected the correct target exemplar, they got both exemplar and category information correct. However, one foil used was a different exemplar of the same category as the target. If this foil was selected, the participant got category information correct but exemplar information incorrect. The other two foils were two exemplars of a different category, if the participant selected either of these foils, they got both category and exemplar information incorrect. All foils were exemplars presented in other sequences from the same study block, controlling for overall stimulus familiarity at test.

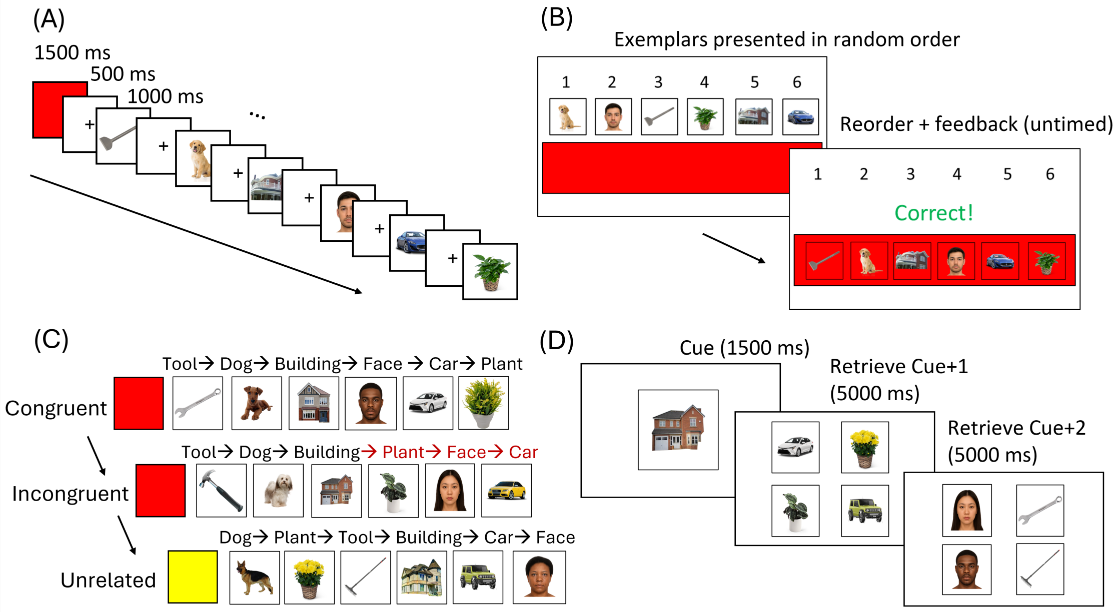

Figure 2.1: Schematic of the experimental procedure. (A) An example of a study trial. Participants would see a context colour telling them what schema they were in, and the associated exemplars would appear one at a time for a second each. Participants were required to make an animacy judgement on each image while it was on the screen. (B) A test trial during pre-learning. After seeing a single sequence as in panel A, participants would see the same exemplars in a random order and would use the 1-6 keys to reorder them back into the correct order. Participants received feedback on their accuracy. (C) A study block in the main temporal order memory task. Participants saw three sequences presented in a row in the same manner as in panel A. The figure shows a congruent and an incongruent sequence both of which are from the red schema. However, it was also possible for both of these to be congruent or both to be incongruent. Likewise, both could be from the blue schema, or one could be from the red schema and one could be from the blue schema. A block always contained a single unrelated trial with yellow as the context colour. (D) A test trial in the main temporal order memory task. Participants were cued with an exemplar from one of the three previous sequences and were required to retrieve Cue+1 and Cue+2 in a four-alternative forced-choice task. The sequence tested in panel D is the incongruent sequence presented in panel C.

2.2.1.5 Analysis approach

Analyses were conducted in R 4.3.3 (R Core Team, 2023) using the *Buildmer* package (Voeten, 2020). Alpha level was set at 0.05. Buildmer was provided with a maximal model including all fixed and random effects. From this maximal model, Buildmer determined the maximal model that still converged and performed stepwise elimination of random effects to determine the most parsimonious model. Analysis of accuracy for both the encoding and retrieval tasks used logistic mixed-effects (GLME) models. Response time analysis for the encoding task used a linear mixed-effects (LME) model. The fixed effects of Congruency and Test were coded using treatment contrasts. The contrast of Congruency used the Unrelated condition as the baseline. The contrast of Test used Cue and Cue+1 as the baselines in the encoding and retrieval analyses respectively.

Brehm & Alday (2022) have highlighted that contrast coding in models such as the one described above are often misinterpreted. When using treatment contrasts, the effect of each factor (Congruency or Test) is assessed at the baseline level of the other factor. Therefore, in the present model, the comparisons of Congruency (Congruent vs. Unrelated; Incongruent vs. Unrelated) will be performed at the baseline level of Test; those baselines being Cue and Cue+1 for the encoding and retrieval analyses respectively. Likewise, the comparison of Test (encoding: Cue vs. Cue+1; retrieval: Cue+1 vs. Cue+2) will be assessed at the baseline of Congruency, i.e., the Unrelated condition. Using treatment contrasts in the present study is crucial as it explicitly sets the Unrelated condition as the control, as intended by the experimental design. However, it also means that our main model leaves some important comparisons unaddressed.

At encoding, the main model does not provide any information concerning how performance differs within each Congruency condition between Cue and Cue+1. Likewise, at retrieval, it does not provide information concerning Congruency effects at Cue+2 nor when collapsing across the levels of Test. As such, I broke down the models using the R package *emmeans*. At encoding, I first investigated all three pairwise comparisons of the Congruency factor at Cue and Cue+1 separately. This allowed me to assess how performance in the Congruent and Incongruent conditions differed from the Unrelated condition (and each other) both before and after the violation in the Incongruent condition. Second, I compared performance in each Congruency condition across Cue and Cue+1 to assess whether the violation at Cue+1 influenced performance in the Incongruent condition specifically. At retrieval, I first investigated all three comparisons of the Congruency factor while collapsing across Test. I then investigated these comparisons at Cue+1 and Cue+2 separately to assess any interactions between Congruency and Test. This allowed me to assess how recall differed across the levels of Congruency. As I conducted all three pairwise comparisons, the p-values reported in the text are all Holm corrected for multiple comparisons. However, the uncorrected outputs of the main mixed-effects models can be found in the tables referenced throughout the text.

By assessing differences in response time and accuracy at encoding I aimed to provide evidence that participants were using the schemas to make predictions about what is likely to happen next in a sequence, and that these ongoing predictions were followed by enhanced memory in the Congruent and Incongruent conditions. Specifically, I made the following predictions:

1. (H1) At encoding, animacy judgements would be faster and more accurate in the Congruent condition relative to the Unrelated condition.
2. (H2) At encoding, animacy judgements would be less accurate and slower in the Incongruent condition on Cue+1 trials as compared to Cue trials.
3. (H3) At retrieval, both category and exemplar-level performance would be enhanced in the Congruent and Incongruent conditions relative to the Unrelated condition at both Cue+1 and Cue+2.

Throughout the results section, analyses addressing the above hypotheses are accompanied by other complementary analyses. As such, the statistics that address the above predictions will be accompanied with a note (e.g., H1) highlighting that those analyses directly test the predictions made here. Experiment 1 was not pre-registered, however, Experiments 2-4 were.

2.2.2 Results

2.2.2.1 Pre-learning

Participants learned the sequence schemas across two sessions and were required to reach criterion in both sessions. I ran an analysis on trials-to-criterion in each session to assess whether participants were retaining knowledge from the first pre-learning session. A paired t-test revealed that participants required more trials to reach criterion in session 1 ($M =$ 49.9, $SD =$ 26.56) than in session 2 ($M =$ 19.07, $SD =$ 8.2), $t(29) =$ 5.8, p <.001 , $d =$ 1.57. This suggests that, coming into session 2, participants were retaining knowledge of the schemas from the first pre-learning session.

2.2.2.2 Encoding

To be included in the encoding analyses participants were required to be significantly above chance (50%) on the animacy judgement task as determined by a one-tailed binomial test. The final sample size for the encoding analysis was $n =$ 26.

Mean accuracy and response times across the Congruency and Test manipulations can be seen in Table 2.1. A full summary of the mixed-effects model at encoding in Experiment 1 can been seen in Table 2.2. The results of the encoding analysis are plotted in Figure 2.2.

Table 2.1: Mean accuracy and response times at encoding between Cue and Cue+1 across the three Congruency conditions in Experiment 1. Standard deviations are shown in parentheses.

Metric	Congruency	Cue	Cue+1
Accuracy (%)	Unrelated	80.77 (12.29)	80.15 (12.11)
	Congruent	85.65 (10.63)	88.91 (10.51)
	Incongruent	87.82 (11.18)	66.99 (17.91)
Response Time (ms)	Unrelated	627.36 (64.89)	637.14 (68.97)
	Congruent	570.51 (88.97)	591.64 (87.23)
	Incongruent	581.74 (85.51)	666.28 (71.67)

2.2.2.2.1 Accuracy

I first assessed whether accuracy differed as a function of Congruency on Cue trials (where the categories were still congruent on both the Congruent and Incongruent trials). Accuracy was significantly greater in the Congruent ($\beta = 0.38$, *SE* = 0.15, $z = 2.47$, $p = .027$; H1) and Incongruent ($\beta = 0.58$, *SE* = 0.16, $z = 3.56$, $p = .001$) conditions on Cue trials as compared to Unrelated Cue trials. This suggests participants had some knowledge of the category (or animacy or motor response) associated with the next trial in the sequence in both the Congruent and Incongruent conditions relative to the Unrelated condition.

I next assessed differences between the Congruency conditions at Cue+1 (after the change had occurred in the Incongruent condition). Accuracy was greater in the Congruent condition as compared to the Unrelated ($\beta = 0.73$, *SE* = 0.16, $z = 4.51$, $p < .001$; H1) and Incongruent conditions ($\beta = 1.47$, *SE* = 0.15, $z = 9.54$, $p < .001$). In contrast to the results at Cue, performance in the Incongruent condition at Cue+1 was significantly lower than performance in the Unrelated condition ($\beta = -0.74$, *SE* = 0.14, $z = -5.52$, $p < .001$). This indicates that the incongruency that occurred at Cue+1 may have disrupted any ongoing predictions that were being used to facilitate performance.

Notably, in line with the above comparisons, the main model revealed a significant Congruency by Test interaction for the Incongruent vs. Unrelated comparison ($\beta = -1.32$, *SE* = 0.21, z = -6.27, p <.001) but there was no interaction for the Congruent vs. Unrelated comparison ($\beta = 0.35$, *SE* = 0.22, z = 1.58, p = .114).

I then investigated differences within each Congruency condition across Cue and Cue+1. Here, I found that there was no difference in accuracy between Cue and Cue+1 in the Congruent ($\beta = 0.31$, *SE* = 0.19, z = 1.59, p = .111) or Unrelated condition ($\beta = -0.05$, *SE* = 0.17, z = -0.27, p = .790). However, there was a significant decrease in accuracy from Cue to Cue+1 in the Incongruent condition ($\beta = -1.37$, *SE* = 0.18, z = -7.72, p <.001; H2). Again, this indicates that the violation of the schemas at Cue+1 disrupted performance in the Incongruent condition, potentially due to a disruption of ongoing predictions of upcoming events.

2.2.2.2.2 Response time

Note, upon inspection of the response time data there were a very small number of negative response times. All of these responses occurred within 16 ms of stimulus onset which is within the frame rate of the monitor used (1000 ms / 60 Hz = 16.7 ms). When the screen was flipped, there was up to a 16.7 ms delay before the stimulus actually appeared. Although responses could be registered as soon as the screen was flipped, response times were calculated against the time when the stimulus actually appeared on the screen. These negative response times are therefore a consequence of participants making anticipatory responses, as correct responses were predictable in this paradigm, and the response being registered in this 16 ms window before the stimulus actually appeared. Given that anticipatory responses were expected in this paradigm, I have opted to keep these negative response times in the dataset for completeness.

On cue trials, response times were significantly faster in the Congruent ($\beta = -53.98$, *SE* = 8.92, z = -6.05, p <.001; H1) and Incongruent ($\beta = -45.07$, *SE* = 9.04, z = -4.98, p <.001) conditions as compared to Unrelated Cue trials. This indicates that prior to any violations in the Incongruent condition, knowledge of the sequence schemas was facilitating faster responses in the Congruent and Incongruent conditions.

I next assessed differences between the Congruency conditions at Cue+1 (after the change had occurred in the Incongruent condition). Response times were faster in the Congruent condition as compared to both the Unrelated ($\beta = -46.5$, *SE* = 8.86, z = -5.25, p <.001; H1) and Incongruent conditions ($\beta = -77.16$, *SE* = 9.44, z = -8.17, p <.001). Notably, response times were also slower in the Incongruent condition relative to the Unrelated condition ($\beta = 30.66$, *SE* = 9.73, z = 3.15, p = .002).

Once again, there was also a significant Congruency by Test interaction for the Incongruent vs. Unrelated comparison ($\beta = 75.72$, *SE* = 13.28, z = 5.7, p <.001), however, there was no interaction for the Congruent vs. Unrelated comparison ($\beta = 7.48$, *SE* = 12.57, z = 0.59, p = .552)

Finally, when investigating the differences between Cue and Cue+1 within each congruency condition I found that there was no difference in response times between Cue and Cue+1 in the Unrelated condition ($\beta = 12.27$, *SE* = 9.15, z = 1.34, p = .180). However, response times were slower at Cue+1 as compared to Cue trials for both the Congruent ($\beta = 19.75$, *SE* = 8.62, z = 2.29, p = .022) and Incongruent conditions ($\beta = 87.99$, *SE* = 9.63, z = 9.14, p <.001; H2). It is unclear why response times would be slower in the Congruent condition at Cue+1, however, response times in the Congruent condition remained faster than those in the Unrelated condition indicating that the schemas were still facilitating performance as compared to the baseline. Once again, the increase in response times between Cue and Cue+1 in the Incongruent condition indicates that the incongruency manipulation disrupted performance due to a violation of prior knowledge.

Table 2.2: Summary of the mixed-effects model at encoding for both accuracy and response time in Experiment 1. P-values are uncorrected and therefore may differ from those in the text.

Metric	Fixed effect	Estimate	SE	z/t	p
Accuracy (%)	Intercept	1.57	0.17	9.21	<.001
	Congruent (vs. unrelated)	0.38	0.15	2.47	.014
	Incongruent (vs. unrelated)	0.58	0.16	3.56	<.001
	Test (Cue vs. Cue+1)	-0.05	0.17	-0.27	.790
	Congruent x Test	0.35	0.22	1.58	.114
	Incongruent x Test	-1.32	0.21	-6.27	<.001
Response Time (ms)	Intercept	625.08	15.48	40.39	<.001
	Congruent (vs. unrelated)	-53.98	8.92	-6.05	<.001
	Incongruent (vs. unrelated)	-45.07	9.04	-4.98	<.001
	Test (Cue vs. Cue+1)	12.27	9.15	1.34	.180
	Congruent x Test	7.48	12.57	0.59	.552
	Incongruent x Test	75.72	13.28	5.70	<.001

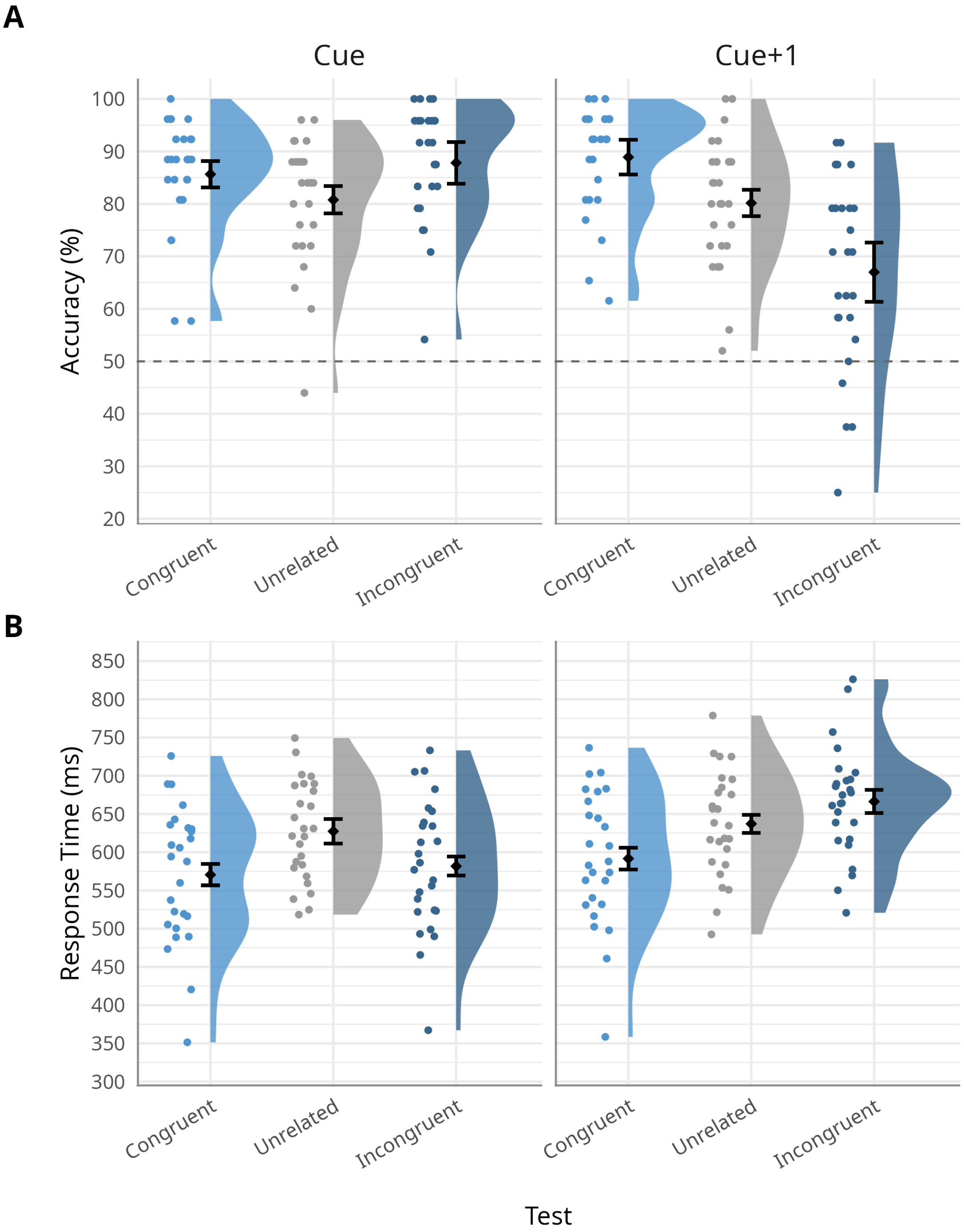

Figure 2.2: (A) Mean accuracy (%) and (B) response time (ms) across the Congruency and Test manipulations at encoding in Experiment 1. Diamonds show sample means. Density and scatter plots show the distribution of participant means. Error bars represent 95% confidence intervals.

2.2.2.3 Retrieval

Mean accuracy across the Congruency and Test manipulations at retrieval can be seen in Table 2.3. A full summary of the mixed-effects model at retrieval can be seen in Table 2.4. The results of the retrieval analysis are plotted in Figure 2.3.

Table 2.3: Mean percentage accuracy at retrieval between Cue+1 and Cue+2 across the three Congruency conditions at the category, exemplar, and conditional exemplar level in Experiment 1. Standard deviations are shown in parentheses.

Level	Congruency	Cue+1	Cue+2
Category	Unrelated	58.53 (6.65)	56.27 (7.92)
	Congruent	64.27 (8.17)	64.67 (5.94)
	Incongruent	60.27 (10.17)	58.67 (8.93)
Exemplar	Unrelated	38.11 (8.74)	34 (9.36)
	Congruent	38.89 (10.52)	40 (6.96)
	Incongruent	36.11 (8.33)	32.22 (9.05)
Conditional Exemplar	Unrelated	69.31 (10.92)	66.78 (8.84)
	Congruent	65.44 (12.81)	63.72 (8.52)
	Incongruent	65.06 (6.53)	58.72 (10.42)

2.2.2.3.1 Category level

To be included in the category-level retrieval analysis participants were required to be significantly above chance (50%) in terms of category-level accuracy collapsed across Congruency and Test manipulations. This resulted in a final sample size of $n = 15$.

Collapsing across Test (Cue+1 and Cue+2), I saw that accuracy was significantly greater in the Congruent condition as compared to both the Unrelated ($\beta = 0.3$, $SE = 0.08$, $z = 3.96$, $p < .001$) and Incongruent conditions ($\beta = 0.21$, $SE = 0.08$, $z = 2.82$, $p = .010$).

There was no difference between the Incongruent and Unrelated conditions ($\beta = 0.09$, *SE* $= 0.07$, $z = 1.15$, $p = .251$).

Despite no evidence for a Congruency by Test interaction (see Table 2.4), I assessed the Congruency effect at Cue+1 and Cue+2 separately for completeness. After correction for multiple comparisons there were no significant effects at Cue+1 at the category level. When investigating the differences across the Congruency conditions at Cue+2 I found that accuracy was higher in the Congruent condition relative to the Unrelated condition ($\beta = 0.35$, *SE* $= 0.11$, $z = 3.33$, $p = .003$; H3) and relative to the Incongruent condition ($\beta = 0.25$, *SE* $= 0.11$, $z = 2.39$, $p = .034$). There was no difference between the Incongruent and Unrelated conditions ($\beta = 0.1$, *SE* $= 0.1$, $z = 0.94$, $p = .347$; H3). Overall, these results suggest the presence of a congruency effect at the category level, however, there was no indication of an effect due to schema incongruency.

2.2.2.3.2 Exemplar level

To be included in the exemplar-level retrieval analysis participants were required to be significantly above chance (25%) in terms of exemplar-level accuracy collapsed across Congruency and Test manipulations. This resulted in a final sample size of $n = 18$.

Collapsing across Test, neither the Congruent ($\beta = 0.15$, *SE* $= 0.07$, $z = 2.12$, $p = .067$) nor the Incongruent condition ($\beta = -0.08$, *SE* $= 0.07$, $z =$ -1.19, $p = .232$) differed from the Unrelated baseline. However, there was a difference between the Congruent and Incongruent conditions, with performance being worse in the Incongruent condition ($\beta = 0.23$, *SE* $= 0.07$, $z = 3.32$, $p = .003$). Assessing Cue+1 and Cue+2 separately, there were no significant effects at Cue+1 at the exemplar level. However, at Cue+2, I found that accuracy was higher in the Congruent condition as compared to the Unrelated condition ($\beta = 0.26$, *SE* $= 0.1$, $z = 2.65$, $p = .016$; H3) and the Incongruent condition ($\beta = 0.34$, *SE* $= 0.1$, $z = 3.45$, $p = .002$). These results provide some preliminary evidence of a congruency effect at the exemplar level. However, given that exemplar and category-level performance are not independent in this design, I next moved on to investigate exemplar-level accuracy while controlling for performance at the category level.

2.2.2.3.3 Conditional exemplar level

In this experiment, exemplar-level performance is conditional on category-level knowledge. That is, chance performance at the exemplar level is 25% if the participant does not know the correct category. However, if the participant knows the correct category, chance performance at the exemplar level increases to 50%. To account for this, I analysed the exemplar-level data while only including trials on which participants selected the correct category. This allowed me to determine whether overall exemplar-level performance, and any schema effects at the exemplar level, were purely driven by category-level accuracy.

To be included in the conditional exemplar-level analysis participants were required to be above chance performance at the exemplar level when only considering trials on which category-level information was accurately recalled (50%). This resulted in a final sample size of $n = 13$.

Collapsing across Test, I observed that accuracy in the Incongruent condition was significantly lower than accuracy in the Unrelated condition ($\beta = -0.27$, $SE = 0.11$, $z = -2.43$, $p = .045$). However, this contrast was not significant when analysing Cue+1 and Cue+2 trials independently. All other comparisons were non-significant. This indicates that after controlling for category-level performance there were no schema effects at the exemplar level.

Table 2.4: Summary of the mixed-effects model at retrieval for the category, exemplar, and conditional exemplar-level analyses in Experiment 1. P-values are uncorrected and therefore may differ from those in the text.

Level	Fixed effect	Estimate	SE	z	p
Category	Intercept	0.35	0.08	4.41	<.001
	Congruent (vs. unrelated)	0.24	0.11	2.28	.023
	Incongruent (vs. unrelated)	0.07	0.11	0.68	.494
	Test (Cue+1 vs. Cue+2)	-0.09	0.10	-0.89	.374
	Congruent x Test	0.11	0.15	0.73	.463
	Incongruent x Test	0.03	0.15	0.18	.859
Exemplar	Intercept	-0.49	0.09	-5.55	<.001
	Congruent (vs. unrelated)	0.03	0.10	0.34	.733
	Incongruent (vs. unrelated)	-0.09	0.10	-0.88	.377
	Test (Cue+1 vs. Cue+2)	-0.18	0.10	-1.83	.068
	Congruent x Test	0.23	0.14	1.64	.100
	Incongruent x Test	0.01	0.14	0.04	.969
Conditional Exemplar	Intercept	0.83	0.12	6.65	<.001
	Congruent (vs. unrelated)	-0.19	0.15	-1.26	.206
	Incongruent (vs. unrelated)	-0.20	0.16	-1.29	.197
	Test (Cue+1 vs. Cue+2)	-0.12	0.16	-0.75	.456
	Congruent x Test	0.06	0.22	0.26	.798
	Incongruent x Test	-0.13	0.22	-0.60	.547

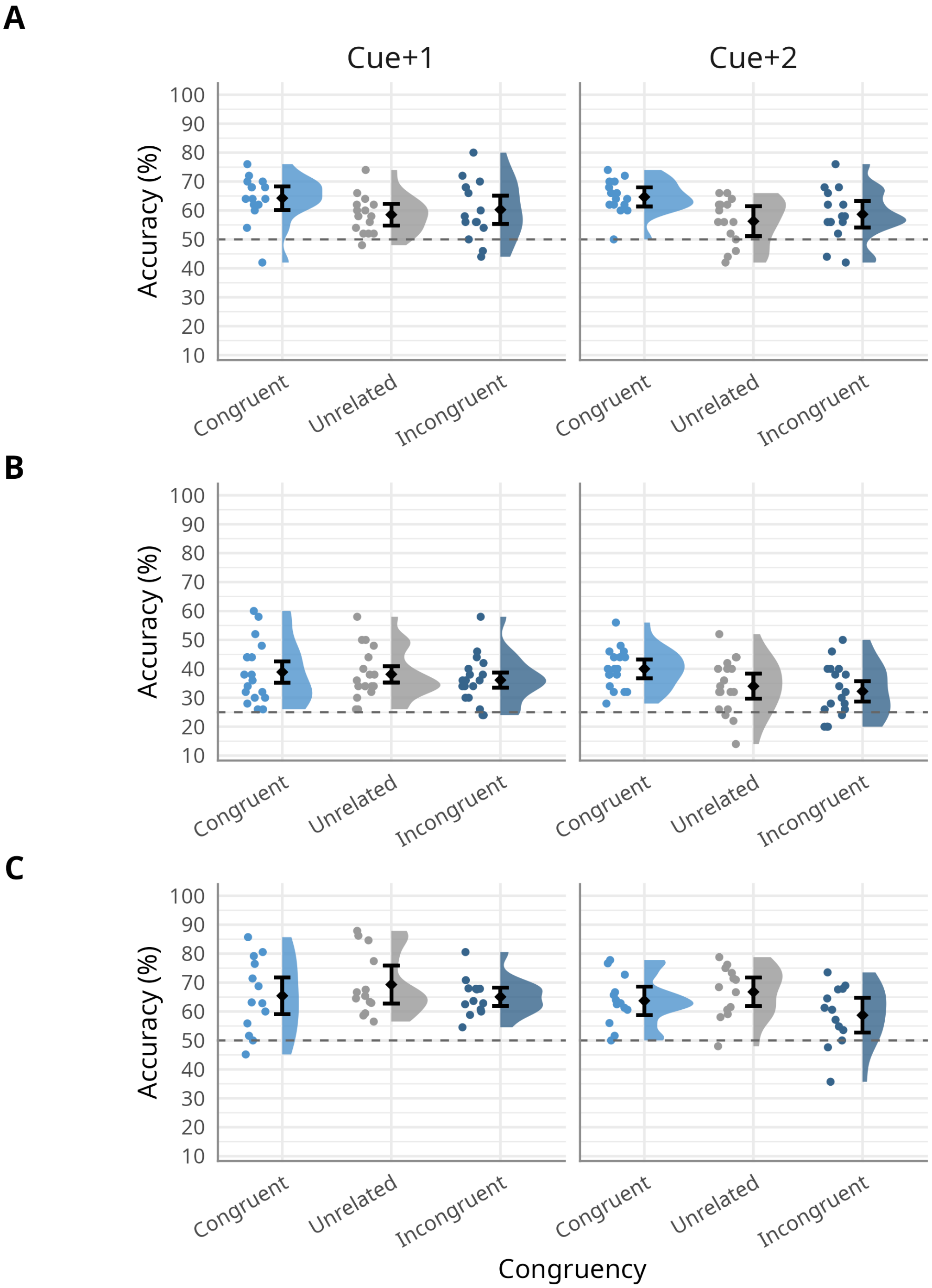

Figure 2.3: Mean accuracy (%) at retrieval across the Congruency and Test manipulations at (A) category level, (B) exemplar level, and (C) conditional exemplar level in Experiment 1. Diamonds show the sample means. Density and scatter plots show the distribution of participant means. Error bars represent 95% confidence intervals.

2.2.3 Discussion

In Experiment 1, I provided evidence that participants are able to use their prior knowledge of sequence structure to make predictions regarding upcoming items. Notably, before any schema violations, performance in the animacy categorisation task at encoding was enhanced in both the Congruent and Incongruent conditions relative to the Unrelated baseline. However, after the schema violation, performance was impeded in the Incongruent condition resulting in accuracy and response times that were impaired relative to the Unrelated baseline condition. Crucially, at test, these apparent predictions were followed by evidence of a schema congruency effect. At the category level, retrieval accuracy was greater in the Congruent condition as compared to the Unrelated condition. However, in contrast to the prior literature on this topic, I saw no evidence of an incongruency advantage in Experiment 1 and there was little evidence of any schema effects at the exemplar level.

What might explain the lack of an incongruency effect? It is thought that incongruency advantages arise due to prediction errors triggered by unexpected events. However, in the present study, as a participant sees more and more Incongruent trials, schema incongruency may become somewhat expected, weakening any prediction errors. This problem may be exacerbated by the fact that Experiment 1 contained an equal number of Congruent and Incongruent trials, meaning that schema incongruency was just as common as schema congruency. In the next experiment, I therefore tried to address this issue by introducing some controls that were aimed at increasing the unexpectedness of the Incongruent trials. First, I introduced extra congruent filler blocks to reinforce the congruent schemas throughout the experiment and to make schema congruency more expected than schema incongruency. I also implemented pseudo-randomisation of the blocks to ensure that large chunks of Incongruent trials could not occur back-to-back, as large chunks of incongruency may reinforce the idea that the schemas cannot be used to make accurate predictions. Finally, in Experiment 1, it was possible for the very first blocks of the experiment to include Incongruent trials (and could potentially include only Incongruent and Unrelated sequences). Exposing participants to Incongruent trials early in the experiment could potentially set a precedent for the rest of the task, indicating to participants that the sequences do not conform to the pre-learned structures. Therefore, in Experiment 2, I forced the first three blocks to only include Congruent sequences in order to reinforce the idea that the sequences in the main temporal order memory task conformed to the pre-learned structures.

It is also noteworthy that Experiment 1 saw a large number of exclusions due to poor overall performance. To be included in any of the analyses, participants were required to perform at an above chance level as determined by a one-tailed binomial test on their trial-level data. Given the large number of trials per participant (150), this is by no means a conservative criterion, and the lack of participants achieving this level of performance suggests the task may be too difficult in general. As such, Experiment 2 aimed to improve overall performance by reducing the length of the sequences from six items to four.

Overall, Experiment 2 had three aims: (1) replicate the schema congruency effects observed in Experiment 1 in a pre-registered design (note, Experiment 1 was not pre-registered), (2) try to induce an incongruency advantage by controlling the balance of Congruent and Incongruent trials throughout the experiment, and (3) improve overall performance and reduce the number of exclusions due to low accuracy by reducing the sequence length from six items to four.

2.3 Behavioural Experiment 2

2.3.1 Methods

2.3.1.1 Participants

Based on the data collected from $n = 16$ participants who were above chance at the category level in Experiment 1, I ran a power analysis for generalised mixed-effects models using the R package *SIMR*. From this, a sample size of $N = 45$ was determined, which gives 99% power to detect a congruency effect assuming an effect size of at least $d = 0.28$ and 80% power to detect an incongruency effect assuming an effect size of at least $d = 0.12$. Thus, Experiment 2 had at least 80% power to detect a "small" effect ($d < 0.20$). For consistency across experiments, all further behavioural experiments aimed for a sample size of $N =$ 45.

Fifty-six participants were recruited via SONA systems. Two participants were excluded for not being above chance at encoding or retrieval as determined by a one-tailed binomial test. This left a final sample of $N = 54$ (44 females, 8 males, 2 prefer not to say; $M \pm SD$ age in years $= 19.74 \pm 1.81$). All participants were fluent English speakers with normal or corrected-to-normal vision. Participants gave informed consent for the experiment, which was approved by the Research Ethics Committee of the Department of Psychology, University of York. Participants were reimbursed for their time with either course credit or a £15 Amazon voucher.

2.3.1.2 Procedure

Experiment 2 aimed to replicate the effects observed in Experiment 1 but had a few adjustments. Sequence length was reduced from six items to four in order to increase overall performance. I also implemented some controls to try and promote the emergence of an incongruency advantage. Experiment 2 included 12 congruent filler blocks. Filler blocks contained either two red and one blue sequence, or two blue and one red sequence, all of which were schema-congruent. Filler blocks were pseudo-randomly spread throughout the task making congruency more common than incongruency across the experiment. Filler blocks were removed from the data prior to analysis. In addition to adding the filler blocks, I also made it so that the first three blocks of the task were guaranteed to only include Congruent and Unrelated sequences. I also pseudorandomised the remaining blocks to prevent blocks containing two Incongruent trials from occurring back-to-back. This ensured that throughout the experiment a participant could not see more than three Incongruent trials before at least one Congruent trial would occur. Finally, during pre-learning, participants were required to get ten sequences in a row correct (instead of six) before reaching criterion and being tested.

Other than the shortening of the sequences from six items to four, and the implementation of the controls outlined above, Experiment 2 had an identical design to Experiment 1. There were 50 sequences for each Congruency condition spread across 50 experimental blocks. There were also 12 congruent filler blocks giving 18 blue and 18 red filler sequences. In total, this gave 62 blocks in Experiment 2. The schemas at the animacy level in Experiment 2 were as follows: (a) *animate-inanimate-animate-inanimate* (b) *inanimate-inanimate-animate-animate*. As before, different participants saw different categories in different positions and which categories a participant saw was counterbalanced across participants.

It is also noteworthy that due to the reduction in sequence length, controls had to be put in place regarding the order of the categories in the Unrelated condition. When using six-item sequences there are 720 possible permutations of the six categories. As such, in Experiment 1, Unrelated sequences were simply the categories presented in a random order. Given the large number of permutations, it was possible, but unlikely that the Unrelated sequence could closely resemble one of the schemas. However, with four-item sequences there are only 24 possible permutations of the categories, meaning that it was highly likely that an Unrelated sequence could closely resemble one of the schemas. To control for this, the Unrelated sequences could only take one of six predetermined orders that were determined to be distinct enough from the schemas such that knowledge of schema structure could not enhance performance in the Unrelated condition. Despite this, participants were still told that the yellow (Unrelated) sequences occurred in a random order.

The shortening of the sequences also had consequences for where I could cue from. In Experiment 1, I used either the second or third item in the sequence as a cue. However, with four-item sequences, it was not possible to cue using the third item as this would not leave enough remaining categories to test both Cue+1 and Cue+2. Therefore, Experiment 2 cued participants with either the first or the second item in a sequence. Critically, this meant that incongruency now occurred either on the second or third item in a sequence.

Finally, in Experiment 1, all foils at test came from different sequences from the same study block. However, given the shortening of the sequences in Experiment 2, the number of exemplars that could be selected from to be used as foils was reduced. This meant that on a minority of trials in Experiment 2, it was possible for a foil exemplar to come from the same sequence as the one being tested.

2.3.1.3 Analysis approach

The analysis approach used in Experiment 2 mirrored that used in Experiment 1 and was pre-registered on the Open Science Framework (Exp 2 pre-registration).

2.3.2 Results

2.3.2.1 Pre-learning

I ran an analysis on trials-to-criterion in each session to assess whether participants were retaining knowledge from the first pre-learning session. A paired t-test revealed that participants required more trials to reach criterion in session 1 ($M = 32.62$, $SD = 17.81$) than in session 2 ($M = 22.75$, $SD = 11.57$), t(54) = 3.49, p <.001, $d = 0.66$. This suggests that participants had retained knowledge of the sequence schemas acquired during session 1.

2.3.2.2 Encoding

To be included in the encoding analyses participants were required to be significantly above chance (50%) on the animacy judgement task. This resulted in a sample size of $n = 51$. Once again, the data were filtered to only include the sequence items that were used as the Cue and Cue+1 at test.

Mean accuracy and response times across the Congruency and Test manipulations can be seen in Table 2.5. A full summary of the mixed-effects model at encoding in Experiment 2 can been seen in Table 2.6. The results of this analysis are plotted in Figure 2.4.

Table 2.5: Mean accuracy and response times at encoding between Cue and Cue+1 across the three Congruency conditions in Experiment 2. Standard deviations are shown in parentheses.

Metric	Congruency	Cue	Cue+1
Accuracy (%)	Unrelated	75.65 (12.42)	83.25 (9.47)
	Congruent	82.51 (11.52)	89.37 (8.28)
	Incongruent	83.14 (10.41)	78.94 (9.82)
Response Time (ms)	Unrelated	714.45 (55.3)	670.05 (62.8)
	Congruent	666.24 (63.19)	615.97 (78.56)
	Incongruent	665.3 (62.61)	658.26 (66.84)

2.3.2.2.1 Accuracy

First, I investigated whether accuracy differed as a function of Congruency on Cue trials (where the categories were still congruent on both Congruent and Incongruent trials). Accuracy was significantly greater on Congruent ($\beta = 0.52$, $SE = 0.09$, $z = 6.01$, $p <.001$; prereg H1) and Incongruent ($\beta = 0.46$, $SE = 0.09$, $z = 5.24$, $p <.001$) Cue trials as compared to Unrelated Cue trials. This suggests that prior to any schema violations, prior knowledge was being used to facilitate performance in the Congruent and Incongruent conditions.

I next assessed differences between the Congruency conditions at Cue+1 (after the change had occurred in the Incongruent condition). Accuracy was significantly greater in the Congruent condition as compared to both the Unrelated ($\beta = 0.6$, $SE = 0.1$, $z = 6.14$, $p <.001$; prereg H1) and Incongruent conditions ($\beta = 0.89$, $SE = 0.1$, $z = 8.73$, $p <.001$). Crucially, accuracy was now lower in the Incongruent condition as compared to the Unrelated condition ($\beta = -0.29$, $SE = 0.09$, $z =$ -3.35, $p <.001$). Together, these results indicate that the violation of the schemas at Cue+1 resulted in less accurate performance in the Incongruent condition.

In line with the comparisons above, there was also a significant Congruency by Test interaction when comparing the Incongruent and Unrelated conditions ($\beta = -0.75$, *SE* = 0.1, *z* = -7.24, *p* <.001), however, there was no Congruency by Test interaction for the Congruent vs. Unrelated comparison ($\beta = 0.08$, *SE* = 0.11, *z* = 0.7, *p* = .486).

I then compared Cue vs. Cue+1 within each Congruency condition. There was a significant difference in all three comparisons. Performance significantly improved between Cue and Cue+1 trials in both the Congruent ($\beta = 0.53$, *SE* = 0.1, *z* = 5.27, *p* <.001) and Unrelated condition ($\beta = 0.45$, *SE* = 0.09, *z* = 5.12, *p* <.001). However, there was a significant decrease in performance between Cue and Cue+1 in the Incongruent condition ($\beta = -0.3$, *SE* = 0.09, *z* = -3.31, *p* <.001; prereg H2). This decrease in performance in the Incongruent condition suggests that the schema violations perturbed ongoing cognitive processing leading to reduced accuracy. However, it is unclear why I observed an increase in performance in the Congruent and Unrelated conditions between Cue and Cue+1. This may be a consequence of shortening the sequences in Experiment 2 and a more detailed discussion of why this occurred is included in the discussion. Overall, this pattern of results is highly consistent with the pattern of results observed in Experiment 1.

2.3.2.2.2 Response time

During Cue trials response times were significantly faster in the Congruent ($\beta = -46.93$, *SE* = 6.23, *z* = -7.53, *p* <.001; prereg H1) and Incongruent ($\beta = -48.76$, *SE* = 5.46, *z* = -8.93, *p* <.001) conditions as compared to the Unrelated trials. Once again this indicates that prior to any schema violation, prior knowledge was facilitating performance in the Congruent and Incongruent conditions.

I next assessed differences between the Congruency conditions at Cue+1 (after the change had occurred in the incongruent condition). Response times were faster in the Congruent condition as compared to both the Unrelated ($\beta = -53.75$, *SE* = 6.07, *z* = -8.86, *p*

<.001; prereg H1) and Incongruent conditions (β = -42.58, *SE* = 4.99, z = -8.53, p <.001). However, in contrast to my findings in the accuracy analysis, response times in the Incongruent condition at Cue+1 were still faster than response times in the Unrelated condition (β = -11.17, *SE* = 5.41, z = -2.06, p = .039). These results indicate that response times were not significantly impacted by the schema violations at Cue+1.

Once again there was a significant Congruency by Test interaction when comparing the Incongruent and Unrelated conditions (β = 37.59, *SE* = 6.74, z = 5.58, p <.001) but no interaction when comparing the Congruent and Unrelated condition (β = -6.82, *SE* = 6.64, z = -1.03, p = .305).

Finally, when comparing response times between Cue and Cue+1 within each Congruency condition I observed a significant increase in response speed in both the Congruent (β = -50.25, *SE* = 5.73, z = -8.76, p <.001) and Unrelated conditions (β = -43.43, *SE* = 5.89, z = -7.38, p <.001). However, in contrast to the accuracy analysis, there was no difference in response times between Cue and Cue+1 in the Incongruent condition (β = -5.84, *SE* = 5.84, z = -1, p = .317; prereg H2). Overall, these results may suggest that, in contrast to Experiment 1, there was no effect of incongruency on response times. However, I address a potential explanation for this discrepancy in an exploratory analysis below.

Table 2.6: Summary of the mixed-effects model at encoding for both accuracy and response time in Experiment 2. P-values are uncorrected and therefore may differ from those in the text.

Metric	Fixed effect	Estimate	SE	z/t	p
Accuracy (%)	Intercept	1.24	0.11	11.62	<.001
	Congruent (vs. unrelated)	0.52	0.09	6.01	<.001
	Incongruent (vs. unrelated)	0.46	0.09	5.24	<.001
	Test (Cue vs. Cue+1)	0.45	0.09	5.12	<.001
	Congruent x Test	0.08	0.11	0.70	.486
	Incongruent x Test	-0.75	0.10	-7.24	<.001
Response Time (ms)	Intercept	713.29	7.68	92.83	<.001
	Congruent (vs. unrelated)	-46.93	6.23	-7.53	<.001
	Incongruent (vs. unrelated)	-48.76	5.46	-8.93	<.001
	Test (Cue vs. Cue+1)	-43.43	5.89	-7.38	<.001
	Congruent x Test	-6.82	6.64	-1.03	.305
	Incongruent x Test	37.59	6.74	5.58	<.001

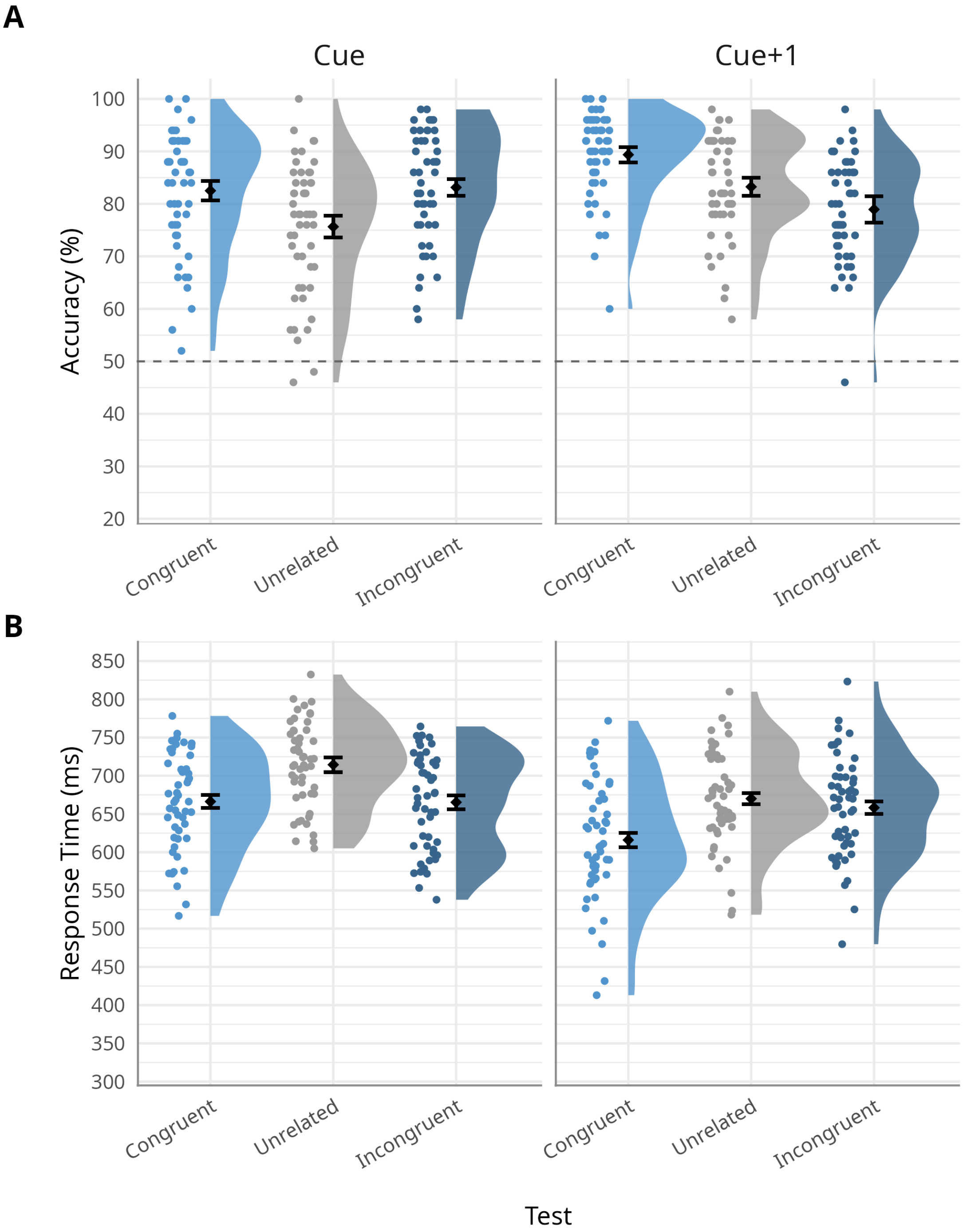

Figure 2.4: (A) Mean accuracy (%) and (B) response time (ms) across the Congruency and Test manipulations at encoding in Experiment 2. Diamonds show sample means. Density and scatter plots show the distribution of participant means. Error bars represent 95% confidence intervals.

2.3.2.2.3 Exploratory encoding analysis

In Experiment 2, I saw that incongruency did not lead to an increase in response times between Cue and Cue+1. This finding is not in line with my prediction that incongruency should induce slower response times, nor the results from Experiment 1 where Incongruent trials led to slower responses. This prediction is based on the notion that participants should be using their prior knowledge to make predictions regarding upcoming categories. However, it is important to highlight that participants are not making motor responses to specific categories, but to the animacy (animate vs. inanimate) of those categories. In Experiment 2, it was possible for an Incongruent Cue+1 trial to be incongruent at the category and animacy level (e.g., *building* to *plant*). However, some trials were only incongruent at the category level (e.g., *building* to *tool*). In this case, although the category differed, the motor response required would be the same as that required by a Congruent sequence. Therefore, if predictions are primarily occurring at the animacy/motor response level, this may account for the lack of an increase in response times between Cue and Cue+1 in the Incongruent condition. To investigate this, I ran a 2x2 ANOVA on the Incongruent data with Test (Cue vs. Cue+1) and the type of Incongruency (Category+animacy vs. Category-only) as predictors.

Accuracy

When considering the accuracy data, there was a main effect of Incongruency, $F(1, 50) = 54.37$, $p < .001$, $\eta^2 = 0.19$ and a significant Test by Incongruency interaction, $F(1, 50) = 59.59$, $p < .001$, $\eta^2 = 0.15$. However, there was no main effect of Test, $F(1, 50) = 0.77$, $p = .384$, $\eta^2 < .01$.

I then ran pairwise t-tests to investigate the differences between Cue and Cue+1 within each Incongruency type. When incongruency occurred at both the category and animacy levels, accuracy decreased at Cue+1 ($M = 71.9\%$, $SD = 14.15\%$) relative to at Cue ($M = 82.59\%$, $SD = 12.15\%$), $t(50) = -5.14$, $p < .001$, $d = -0.81$. However, when incongruency only occurred at the category level, accuracy increased at Cue+1 (M = 92.62%, SD = 8.14%) relative to at Cue (M = 84.2%, SD = 10.88%), t(50) = 5.81, p <.001, d = 0.88.

Response time

When considering the response time data, there was a main effect of Test, F(1, 50) = 5.24, p = .026, $\eta^2 = 0.01$, a main effect of Incongruency, F(1, 50) = 56.43, p <.001, $\eta^2 = 0.13$, and a significant Test x Incongruency interaction F(1, 50) = 76.23, p <.001, $\eta^2 = 0.09$.

When breaking the model down, I found that when Cue+1 was incongruent at both the category and animacy level, response times were significantly slower at Cue+1 (M = 699.92 ms, SD = 57.05 ms) than at Cue (M = 670.12 ms, SD = 69.57 ms), t(50) = 3.58, p <.001, d = 0.47. Moreover, response times at Cue+1 (M = 598.75 ms, SD = 90.23 ms) were faster than at Cue (M = 658.01 ms, SD = 71.39 ms) when Cue+1 was incongruent at only the category level, t(50) = -7.31, p <.001, d = -0.73.

Overall, these findings are in line with the interpretation that predictions in the encoding task are primarily occurring at the animacy and/or motoric level and explain why I failed to see slower response times at Cue+1 in the main encoding analysis.

2.3.2.3 Retrieval

Mean accuracy across the Congruency and Test manipulations at retrieval in Experiment 2 can be seen in Table 2.7. A full summary of the mixed-effects model at retrieval can be seen in Table 2.8. The results of this analysis are plotted in Figure 2.5.

Table 2.7: Mean percentage accuracy at retrieval between Cue+1 and Cue+2 across the three Congruency conditions at the category, exemplar, and conditional exemplar level in Experiment 2. Standard deviations are shown in parentheses.

Level	Congruency	Cue+1	Cue+2
Category	Unrelated	63.43 (10.72)	66.94 (8.64)
	Congruent	67.35 (8.34)	68.53 (9.23)
	Incongruent	62.2 (11.11)	65.35 (9.49)
Exemplar	Unrelated	44.98 (13.97)	45.47 (13.7)
	Congruent	47.59 (12.53)	46.12 (11.88)
	Incongruent	45.55 (14.08)	43.18 (13.69)
Conditional Exemplar	Unrelated	73.74 (11.23)	70.92 (12.51)
	Congruent	73.83 (11.33)	70.02 (10.62)
	Incongruent	75.09 (12.11)	69.35 (12.5)

2.3.2.3.1 Category level

To be included in the category-level retrieval analysis participants were required to be significantly above chance (50%) in terms of category-level accuracy collapsed across Congruency and Test. The final sample for the category analysis was $n = 49$. Note that this sample size is greater than the pre-registered $n = 45$. These extra participants were already recruited when $n = 45$ was reached due to the multi-day nature of the study and uncertainty as to how many participants would be above chance.

Collapsing across Test, I saw that accuracy at the category level was greater in the Congruent condition relative to the Unrelated ($\beta = 0.12$, $SE = 0.05$, $z = 2.46$, $p = .028$) and Incongruent conditions ($\beta = 0.18$, $SE = 0.05$, $z = 3.34$, $p = .002$), however, there was no difference between the Incongruent and Unrelated condition ($\beta = -0.06$, $SE = 0.05$, $z =$ -1.3, $p = .193$).

Despite there being no evidence of a Congruency by Test interaction, for completeness I investigated effects at Cue+1 and Cue+2 separately. At Cue+1, category-level accuracy was greater on Congruent trials than Unrelated trials ($\beta = 0.18$, *SE* = 0.07, $z = 2.66$, $p =$ .016; prereg H3) and Incongruent trials ($\beta = 0.23$, *SE* = 0.07, $z = 3.25$, $p = .003$). However, there was no difference between the Incongruent and Unrelated condition ($\beta = -0.05$, *SE* = 0.06, z = -0.81, p = .416; prereg H3). At Cue+2, there were no significant differences. These results indicate the presence of a schema congruency effect at the category level, however, there appears to be no effect of incongruency.

2.3.2.3.2 Exemplar level

To be included in the exemplar-level retrieval analysis participants were required to be significantly above chance (25%) in terms of exemplar-level accuracy collapsed across Congruency and Test manipulations. This resulted in a sample size of $n = 49$

When investigating effects collapsed across Test, accuracy was greater in the Congruent condition as compared to the Incongruent condition ($\beta = 0.11$, *SE* = 0.05, $z = 2.3$, $p =$ .064). However, no other comparisons were significant. Moreover, there were no significant effects at the exemplar level when Cue+1 and Cue+2 were analysed independently. Overall, these results suggest that there were no schema effects at the exemplar level.

2.3.2.3.3 Conditional exemplar level

To be included in the conditional exemplar level analysis participants were required to be above chance performance at the exemplar level when only considering trials on which category-level information was accurately recalled (50%). This resulted in sample size of $n = 39$. There were no significant effects in any of the conditional exemplar-level analyses with $p > .15$ in all cases after correction for multiple comparisons.

Table 2.8: Summary of the mixed-effects model at retrieval for the category, exemplar and conditional exemplar-level analyses in Experiment 2. P-values are uncorrected and therefore may differ from those in the text.

Level	Fixed effect	Estimate	SE	z	p
Category	Intercept	0.57	0.06	9.02	<.001
	Congruent (vs. unrelated)	0.18	0.07	2.66	.008
	Incongruent (vs. unrelated)	-0.05	0.06	-0.81	.416
	Test (Cue+1 vs. Cue+2)	0.16	0.06	2.61	.009
	Congruent x Test	-0.10	0.09	-1.19	.235
	Incongruent x Test	-0.02	0.09	-0.21	.831
Exemplar	Intercept	-0.21	0.08	-2.49	.013
	Congruent (vs. unrelated)	0.11	0.06	1.77	.076
	Incongruent (vs. unrelated)	0.02	0.06	0.39	.698
	Test (Cue+1 vs. Cue+2)	0.02	0.06	0.36	.722
	Congruent x Test	-0.08	0.08	-0.99	.321
	Incongruent x Test	-0.12	0.08	-1.47	.141
Conditional Exemplar	Intercept	1.08	0.10	10.68	<.001
	Congruent (vs. unrelated)	-0.14	0.09	-1.54	.123
	Incongruent (vs. unrelated)	0.02	0.09	0.21	.837
	Test (Cue+1 vs. Cue+2)	0.09	0.09	0.99	.321
	Congruent x Test	-0.07	0.13	-0.53	.599
	Incongruent x Test	-0.17	0.13	-1.36	.175

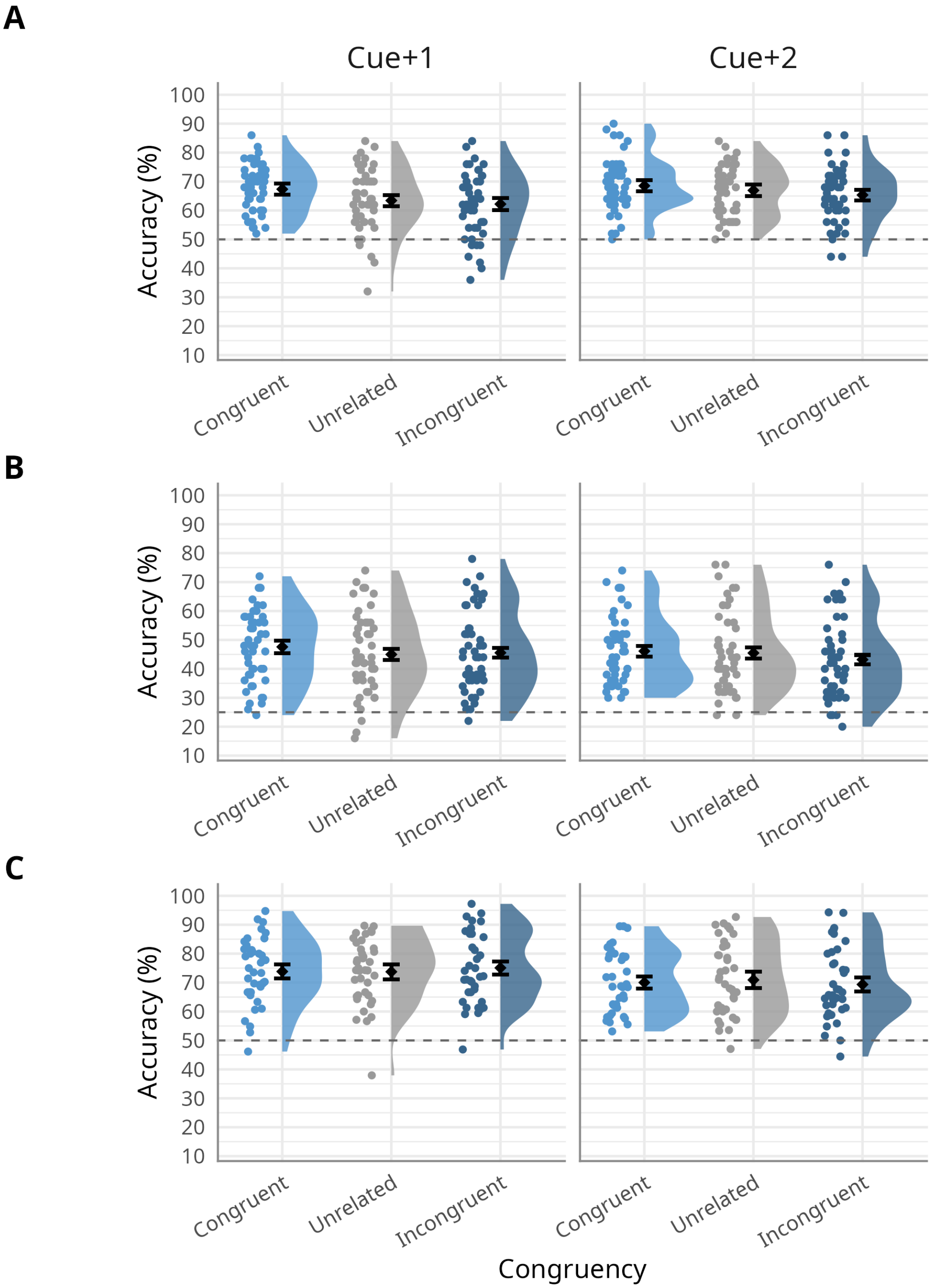

Figure 2.5: Mean accuracy (%) at retrieval across the Congruency and Test manipulations at (A) category level, (B) exemplar level, and (C) conditional exemplar level in Experiment 2. Diamonds show the sample means. Density and scatter plots show the distribution of participant means. Error bars represent 95% confidence intervals.

2.3.2.3.4 Exploratory retrieval analysis

The results at retrieval in Experiment 2 suggest the presence of a congruency advantage at the category but not exemplar level. Although this does indeed indicate that the schemas are modulating performance in some way, it is not clear whether this modulation is occurring at encoding or retrieval. Given that the schemas concern the order of the categories, it is possible to use schema knowledge to aid performance at the point of retrieval. For example, if I know that *building* predicts *dog*, and I am cued with a *building*, I can use my knowledge of the schemas at retrieval to know that I should pick an image of a *dog*. This could produce the observed congruency advantage at the category level without any need for encoding-based schema processes.

I therefore considered ways to assess whether the observed congruency advantage was driven by encoding-based or retrieval-based processes. One way of assessing whether schema knowledge is being used at retrieval is to investigate how the foils used at test modulate performance. It is possible for a foil being used to test a sequence from one schema (schema A) to match what a sequence from the other schema (schema B, that is not being tested) would predict. For example, the schemas could be: (schema A) *face-building-dog-car*, (schema B) *building-car-face-dog*. Assuming we were testing schema A, the participant could be cued with *building* and required to recall *dog*. However, it is possible for the foil category in this case to be *car*, which is what schema B would predict. If participants were using the schemas to facilitate performance at the point of retrieval, we might expect foils that match what the other schema would predict to modulate performance in some way (if participants are not fully aware of which schema is being tested). For example, when testing a Congruent trial, if the foil category matches what the opposite schema would predict, this may negatively impact performance as both the target and the foil category are potentially correct responses based on the schemas.

To test this, I ran an exploratory analysis on category-level accuracy at Cue+1 at each level of Congruency. For each trial, I took the Cue used and determined whether the foil category at Cue+1 matched what the opposite schema would predict. For example, if a participant was being tested on a sequence that used the red schema, the opposite schema would be the blue schema. I then used a t-test to determine whether accuracy could be predicted depending on whether the foil matched what the opposite schema would predict. The results of these analyses are plotted in Figure 2.6.

In the Congruent condition I found that when a foil matched what the opposite schema would predict (M = 66.53%, SD = 12.65%) this did not significantly impact performance as compared to when the foil was not a prediction of the opposite schema (M = 67.69%, SD = 10.38%), $t(48)$ = -0.51, p = .614, d = -0.1. Likewise, I found no difference in accuracy in the Incongruent condition when the foil did (M = 66.88%, SD = 22.07%) or did not match (M = 61.51%, SD = 11.2%) what the opposite schema would predict $t(48)$ = 1.83, p = .074, d = 0.31. However, in the Unrelated condition I found that performance was modulated by the foil used. Specifically, I saw that accuracy was decreased when a foil matched (M = 61.77%, SD = 11.76%) what either the red or blue schema would predict relative to when the foil did not match (M = 65.77%, SD = 13.59%) the prediction of either of these schemas $t(48)$ = -2.1, p = .041, d = -0.31.

Overall, this analysis suggests that, at least in the Congruent and Incongruent conditions, the nature of the foils used at test did not tend to impact performance. This finding does not clearly align with a model that suggests the schema effects in this study are purely driven by retrieval-based mechanisms. I discuss these results in more detail in the discussion below.

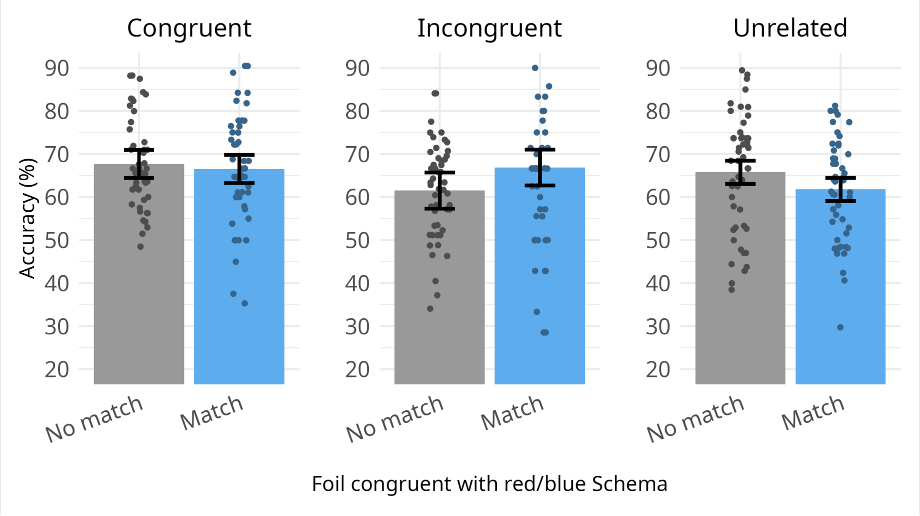

Figure 2.6: Accuracy at the category level on Cue+1 trials broken down by whether the foil category matched what would be predicted by a schema that was not being tested on that trial. Bars show the sample means. Scatter plots show the distribution of participant means. Error bars represent 95% confidence intervals.

2.3.2.4 Experiment 2 Discussion

Experiment 2 successfully replicated the main findings from Experiment 1. Participants showed enhanced reaction times and accuracy at encoding in both the Congruent and Incongruent conditions relative to the Unrelated baseline at Cue, before any violation occurred in the Incongruent condition. At Cue+1, accuracy decreased in the Incongruent condition. However, in contrast to Experiment 1, there was no difference in response time between Cue and Cue+1 in the Incongruent condition.

At retrieval, the congruency advantage at the category level was present at Cue+1 suggesting that this congruency advantage is a robust phenomenon. Notably, as in Experiment 1, there was no effect of incongruency on memory retention. This indicates that the lack of an incongruency advantage in Experiment 1 was not due to poorer overall performance due to the six, rather than four-item sequences. In addition, this suggests that the relative frequency of Congruent and Incongruent sequences in Experiment 1 (equal numbers of each) was not responsible for the lack of an incongruency effect. Experiment 2 contained congruent filler blocks, making congruency more common than incongruency, and an incongruency effect was still not observed.

Exploratory analyses were conducted on both the encoding and retrieval data. These indicated that predictions at encoding were primarily occurring at the animacy, not category level, and that retrieval-based strategies may not fully account for the congruency advantage observed at Cue+1. The implications of these analyses are discussed in more detail in the discussion section below.

2.3.3 Discussion

Across two experiments I investigated whether schema effects could be observed in a temporal order memory paradigm in which a core feature of the schema was the chronology of events. After participants were trained on the sequence schemas, I found that knowledge of these schemas facilitated accuracy and response times during encoding. Moreover, the violation of these schemas in the Incongruent condition at encoding led to a reduction in accuracy and increases in response times. Together these results suggest that the schemas were indeed being used to make predictions of upcoming events. At retrieval, I observed schema congruency effects for the temporal order of the categories in a sequence. However, there was no effect of congruency at the exemplar level and there was no effect of incongruency in any of the retrieval analyses. Existing neurocognitive models of schema effects in memory suggest that schema effects emerge due to encoding-based processes. However, the present results may be more in line with the use of schemas at retrieval instead.

2.3.3.1 Interpretation of encoding results

In the encoding analyses I set out to investigate whether I could index behavioural evidence of ongoing temporal predictions during sequence encoding. Overall, the results support this hypothesis suggesting that prior knowledge can be used to predict upcoming events in this paradigm. Moreover, these predictions were followed by enhanced recollection for the predictable schema-congruent sequences. This is in line with a model that suggests schema effects emerge due to the modulation of encoding-based processes via ongoing predictions. However, it must be noted that although these predictions did occur prior to a boost in memory at retrieval, this study does not provide direct evidence that the congruency advantage was necessarily induced by the affirmation of these predictions. Notably, I also observed that accuracy and response times were degraded in the Incongruent condition at Cue+1 relative to during Cue trials. Again, this is in line with the notion that participants were making predictions of upcoming items in a sequence and that the violation of these predictions impeded performance. However, contrary to what was observed in the Congruent condition, there was no evidence of a memory advantage in the Incongruent condition at retrieval.

This research provides evidence that predictability at encoding can precede enhanced memory for schema-congruent events. However, an important consideration here is at what level these predictions are occurring. The schemas used in the present study contained information regarding the order of the sequences at both the category level (e.g., *face-building-dog-car*) and animacy level (e.g., *animate-inanimate-animate-inanimate*). Importantly, predictions at encoding were assessed based on judgements at the animacy level, raising the question as to whether predictions were being made with regards to the categories as well. The exploratory analysis on this suggests that impairments in performance at Cue+1 in the Incongruent condition were a consequence of incongruency at the animacy level and provided little evidence that incongruency at the category level impacted performance. This indicates that predictions may primarily be occurring at the animacy level with motor response options being made before any semantic assessment of the category. Nevertheless, this analysis only concerned performance between Cue and Cue+1 in the Incongruent condition and thus it cannot be ruled out that predictions at the category level were occurring, especially under schema-congruent circumstances. Further research is required to fully establish whether ongoing neural predictions are taking place at the category level, a question I aim to address with the use of neuroimaging in a later chapter.

In Experiment 2, I also observed an increase in accuracy and a reduction in response times between Cue and Cue+1 in the Congruent and Unrelated conditions. There is no clear theoretical explanation for why this would occur, however, this may be a consequence of the shortening of the sequences in Experiment 2. All sequences in Experiment 2 consisted of two animate and two inanimate categories. Hence, once a participant knows the animacy of the first category, the animacy of the following categories becomes somewhat predictable. For example, if the first category is animate, then there is a two thirds chance that the next category will be inanimate (as the following three categories have to consist of one animate and two inanimate categories). Therefore, even in sequences where the order of the categories is unpredictable (i.e., the Unrelated condition), the animacy of those categories can become more and more predictable as a sequence progresses. Importantly, despite this anomalous finding, overall performance in the Congruent condition was still higher than in the Unrelated condition, and incongruency still led to a reduction in accuracy between Cue and Cue+1. Therefore, this increase between Cue and Cue+1 in the Congruent and Unrelated conditions does not impact our overall conclusions that the schemas could be used to make predictions of upcoming categories.

2.3.3.2 Retrieval vs. encoding-based accounts of schema effects at retrieval.

SLIMM predicts that during the perception of an event, encoding is modulated based on how the experience relates to existing schemas. The observation of a congruency effect at the category level is potentially in line with this model. However, it is notable that schemas may also exert influence at the point of retrieval. That is, although the original instance may not have been encoded sufficiently to facilitate accurate recall at test, the correct response can still be deduced in some cases via recall of the schema itself. Using a restaurant schema as an example, although an individual may not remember specific details regarding a visit to a restaurant, they could still use their schemas to infer that they were seated at their table before they received their food. In the present paradigm, schemas relate to the order of the categories in a given context. Therefore, knowledge of the temporal order of the categories in the schema could be used to enhance category-level performance at the point of retrieval. If you know from your schemas that *face* is reliably followed by *dog*, then when you are cued with a *face*, you can accurately select a *dog* as the subsequent category despite having no memory for the encoding instance. For encoding-based processes to become necessary to explain results in the current study, I would need to observe an incongruency effect at the category level, or a congruency and/or incongruency effect at the exemplar level. However, I did not see any of these effects.

Greve et al. (2019) highlight how the use of so-called generate-and-recognise strategies at retrieval can influence schema effects. If a participant does not know the correct response, they may retrieve a plausible response from memory based on their schemas, and then recognise whether this would be an accurate response in the current context. Crucially, this in turn can impact the emergence of incongruency effects. Under a circumstance in which a participant does not know the correct answer in the Incongruent condition, they may fall back on their knowledge of the schema to try and select the correct category. Notably, in the Incongruent condition, this may actually result in a detriment in performance if the foil used is schema-congruent, as this would result in them selecting the foil category. Indeed, Greve et al. (2019) highlight that the ability to utilise schemas at the point of retrieval may help to explain why many studies observe a congruency advantage but fail to see an effect of incongruency.

However, my exploratory analyses on the foils used at retrieval provide some evidence that retrieval-based mechanisms may not fully account for our observed pattern of results. If the effects were purely driven by retrieval-based mechanisms such as generate-and-recognise, then when a foil is congruent with what another schema would predict, this should have a negative impact on performance, as a participant may be more likely to select that foil (provided they don't know which schema to retrieve). However, I saw that in both the Congruent and Incongruent conditions, the nature of the foils had no significant impact on performance. Interestingly, however, performance was worse in the Unrelated condition when a foil matched what a Congruent schema would predict. Given that this detriment occurred in the Unrelated condition but not in the Congruent or Incongruent conditions, this may indicate generally poorer encoding in the Unrelated condition resulting in a greater reliance on retrieval-based strategies at test. Overall, this analysis provides at least some evidence that retrieval-based strategies were not routinely employed at test; and given that retrieval and encoding-based strategies are not mutually exclusive, this indicates the possibility that encoding-based mechanisms may have contributed to these effects. Nonetheless, this analysis is not conclusive, and retrieval-based mechanisms are still likely to be a major contributing factor to the observed results. Further studies are required to try and establish whether encoding-based mechanisms are contributing to memory in this paradigm. For example, a congruency effect at the exemplar level or an incongruency effect at either level would be difficult to explain via retrieval-based mechanisms. The next section discusses why we may not have observed an incongruency effect and how future studies can try and address this issue.

2.3.3.3 Lack of an incongruency effect

Previous research suggests that schema-incongruent events should be remembered better than events that are unrelated to our existing schemas. However, across two experiments I failed to observe any incongruency advantage. Why might this be? Incongruency effects are believed to be a consequence of prediction errors triggered by unexpected events. Based on mathematical models of learning, the extent of learning is proportional to the size of this prediction error (Rescorla & Wagner, 1972). As such, one simple explanation for the lack of an incongruency effect in this study is a failure to produce a prediction error that was large enough to potentiate learning. The paradigm used here is quite abstract and the reordering of arbitrary categories in this way may simply not be that surprising. This problem may be exacerbated further given that the schemas were pre-trained over only two sessions. Although this provides some period for the knowledge to be consolidated, research suggests that the construction of a unified schema in memory is dependent on slow-wave sleep (Durrant et al., 2011), and it may be that several offline consolidation periods are required before a schema is established enough for its violation to become particularly surprising. However, incongruency effects have been observed in studies that have used shorter consolidation periods than the present study (Greve et al., 2019), indicating that this is unlikely to be the reason why I did not observe one.

It should also be noted that the studies presented in this chapter used overlapping schemas. That is, both schemas used the same categories but presented in different orders. This could potentially lead to interference between the schemas making it more difficult to make accurate predictions of upcoming items. For example, in the red schema, *face* may be followed by *building*, however, in the blue schema, *face* may be followed by *dog*. Therefore, interference may result in poor pattern separation and difficulties in accurately predicting the next item. As such, future studies may benefit from constructing each schema from entirely unique categories in order to reduce any ambiguities as to what each category predicts.

Another potential explanation for the lack of an incongruency effect may come from the nature of the predictions taking place. Davachi & DuBrow (2015) have suggested that when a sequence is learned through repetition, this results in the construction of a predictive code in the hippocampus. When a sequence is encountered numerous times, enhanced hippocampal activity can be observed during sequence perception. This increase in activity has been hypothesised to be indicative of hippocampal pattern completion, the process by which a partial cue from an episode can result in the holistic recollection of the entire event (Horner et al., 2015; Marr, 1971; Nakazawa et al., 2002; Rolls, 2013). Thus, it may be possible that as an individual views a sequence they are predicting the next item in the sequence. For example, if you know from your schemas that *face* is reliably followed by *dog*, you may be predicting *dog* either during the presentation of *face* or during the inter-stimulus interval between the two stimuli. However, it is also possible that when you begin to view a sequence there is holistic reactivation of the entire sequence representation, such that all upcoming elements of the sequence are somewhat pre-activated. Such a model would be in line with models of temporal difference learning that suggest that upcoming events are modeled using a cache, with information concerning all upcoming items being stored in the cache at any given time, but with more proximate items being represented more strongly (Dayan, 1993).

If holistic reactivation of this sort does occur, then although stimuli are presented in a different order in the Incongruent condition, these stimuli are still somewhat predicted and therefore no strong prediction error may be produced. Further studies in the coming chapters will attempt to address this hypothesis by investigating whether different forms of surprise (other than temporal reordering) influence the emergence of incongruency advantages. For example, rather than introducing incongruency effects by reordering existing elements of a sequence, incongruency could be triggered by introducing categories that are completely unexpected in a given context. If incongruency effects could be observed in a such a case this would provide evidence for a model in which predictive representations contain information regarding multiple upcoming categories, rather than individual sequence elements being predicted one at a time.

2.3.3.4 Conclusion

In this chapter, I have successfully developed a temporal order memory task that produces schema congruency effects at the category level. Moreover, these congruency effects are preceded by evidence of ongoing predictions during encoding, indicating that predictions during encoding may contribute to congruency effects at retrieval. Crucially, however, the congruency effects observed at retrieval may not require encoding-based mechanisms to be explained and may be explicable via the use of retrieval-based strategies instead. Given that many models propose that encoding-based mechanisms play an integral role in the generation of schema effects, the next experimental chapter will attempt to find evidence of encoding-based mechanisms in this paradigm. This will be done by (a) using non-overlapping schemas (Experiment 3), and (b) varying the form of violation (Experiment 4).

3 The Influence of Sequential vs. Categorical Violations

3.1 Introduction

In the previous chapter, I was able to demonstrate schema congruency effects for temporal order memory of four-item sequences. Notably, however, I did not see any indication of a schema incongruency effect. Given the strong evidence that incongruency can drive learning, the following chapter aimed at investigating what factors may influence the emergence of incongruency effects in this paradigm. As I see it, there are two predominant reasons why I may not have observed incongruency effects in the previous experiments. First, Experiments 1 and 2 used overlapping schemas. Both schemas contained the same categories presented in different orders. This meant that transitions from one category to the next, although technically predictable, were somewhat ambiguous, and predictions in one schema were susceptible to interference from knowledge of the other schema's structure. This could have inhibited the participants' ability to make reliable predictions, preventing learning via prediction error signals. Second, thus far, I have been working on the assumption that predictions of upcoming categories occur in a chain-like manner. For example, if a sequence is made up of the categories *face-building-dog-car* then participants will predict *face* first, then *building*, then *dog*, and then *car*, in a category-by-category manner as the sequence unfolds. Under this assumption, the change in the temporal order of the categories should conflict with the participant's current prediction, resulting in a prediction error that could enhance learning. However, rather than chaining predictions for each individual category, participants could have a more holistic prediction concerning all categories that are expected throughout the sequence. That is, predictions could be occurring multiple steps into the future, rather than at the level of individual categories, with predictive representations containing information about all upcoming categories. If this were the case, temporal reordering may not trigger a strong prediction error because although the category occurs in an unexpected position it is still included in the current prediction of what items are expected to occur.

Across the two Experiments in this chapter, I investigated these two possibilities. Experiment 3 adopted the same design as Experiment 2, however, the schemas were constructed of unique categories with no overlap in the content of the two schemas. This meant that any given category only ever predicted one other, rather than predicting different things in different contexts. For example, the schemas could be *face-building-dog-tool*, and, *furniture-car-plant-inesct*, such that each category only ever predicts one other category (e.g., *building* always predicts *dog* and nothing else). In theory, this should enhance a participant's ability to predict what is coming next in a sequence, increasing the likelihood of a strong prediction error that could drive learning in the Incongruent condition.

Note, Experiments 3 and 4 are presented in the opposite order to which they were actually conducted. That is, data collection and analysis for Experiment 3 was conducted after Experiment 4 had been completed. I have altered their order here for narrative purposes. This change does not alter the interpretation or implications of the experimental results.

3.2 Behavioural Experiment 3

3.2.1 Methods

3.2.1.1 Participants

Fifty participants were recruited via SONA systems. Two participants were excluded because they did not attend session 2, resulting in a final sample of 48 participants (39 females, 9 males; $M \pm SD$ age in years = 21.85 ± 4.23). All participants were fluent English speakers with normal or corrected-to-normal vision. Participants gave informed consent for the experiment, which was approved by the Research Ethics Committee of the Department of Psychology, University of York. Participants were reimbursed for their time with either course credit or a £15 Amazon voucher.

3.2.1.2 Procedure

The main aim of Experiment 3 was to replicate the findings of Experiment 2 but under conditions in which predictions are unlikely to be perturbed due to interference between overlapping schemas. To achieve this, I implemented the same design as that used in Experiment 2 that used four-item sequences. The key difference in Experiment 3 was that each schema was constructed of unique categories that never occurred in the opposite schema. For example, the schemas for a given participant could be: (red) *face-building-dog-tool*, (blue) *furniture-car-plant-insect*. The animacy structure of the schemas was kept the same as Experiment 2: (red) *animate-inanimate-animate-inanimate*, (blue) *inanimate-inanimate-animate-animate*. Incongruency was still defined as a reordering of the remaining categories after the Cue. The differences in design between Experiments 2 and 3 can be seen in Figure 3.1.

A notable change in Experiment 3 was that the Unrelated sequences could now contain any of the eight categories presented in a pseudo-random order. Participants were told that these sequences were random, however, these sequences were carefully constructed to ensure that they provided the foils needed at test. For example, in a block that contained a red, blue and yellow sequence, the lack of any overlap between the red and blue schemas meant that they could not provide foils for each other at test. If a block contained a red sequence consisting of *face-building-dog-tool* and a blue sequence consisting of *furniture-car-plant-insect,* both of which were using their second item as the Cue, it would be necessary to test participants on their memory for a *dog*, a *tool*, a *plant*, and an *insect*. As such, the Unrelated sequence was forced to include a *dog*, a *tool*, a *plant*, and an *insect* to ensure the necessary foils were available at test.

The new structure of the Unrelated sequences potentially makes them somewhat predictable. As such, I also administered a post-experimental questionnaire asking participants about any strategies they were using to try and remember the sequences and what they thought the experimental manipulation was. A qualitative review of the questionnaires indicated that the majority of participants noticed the incongruency manipulation ($n = 33$) and that the sequences did not always conform to the pre-learned structures. Despite this, there was little indication that any of the participants gained explicit knowledge of predictability in the Unrelated sequences or tried to employ any strategies based on this predictability.

Due to the lack of overlap between the schemas it was also no longer possible to have congruent filler blocks that contained both red and blue sequences (e.g., *red-red-blue*) as there would not be any foils for the sequence that only occurred once. As such, congruent filler blocks in this chapter contained three of only one type of sequence (i.e., *red-red-red* or *blue-blue-blue*). The number of filler blocks was also increased to fourteen, giving 21 red and 21 blue filler sequences in total. This meant Experiment 3 had 64 blocks in total, 50 of which were included in the final analysis.

3.2.1.3 Analysis approach

The analysis approach used in Experiment 3 mirrored that used in Experiment 2 and was pre-registered on the Open Science Framework (Exp 3 pre-registration).

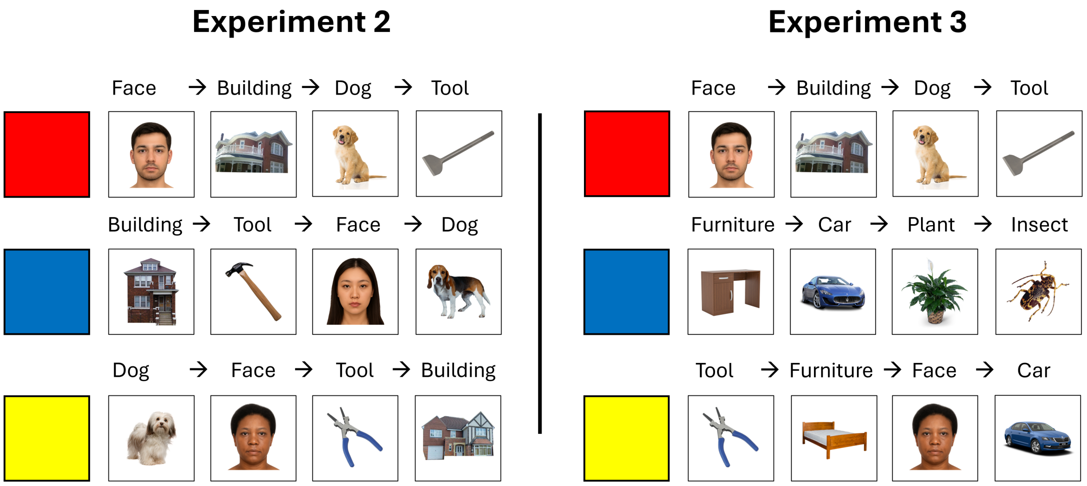

Figure 3.1: The differences in design between Experiments 2 and 3. In Experiment 2, all sequences were created from the same four categories presented in different orders. The red and blue schemas therefore had overlapping categories that differed in temporal order. This could make it harder to make accurate predictions because the same category could predict different things in different contexts. For example, in the red schema, *building* predicts *dog*, however, in the blue schema, *building* predicts *tool*. In Experiment 2, the yellow Unrelated sequences contained the same four categories as the Congruent red and blue sequences but were presented in one of six pre-determined orders that were distinct enough from the schemas such that schema knowledge could not facilitate performance on Unrelated trials. In Experiment 3, the schemas were constructed of unique, non-overlapping categories. As such, each category only ever predicted one other category. In Experiment 3, the Unrelated condition could now contain any of the eight categories.

3.2.2 Results

3.2.2.1 Pre-learning

First, I ran a trials-to-criterion analysis on the pre-learning data to establish whether participants were reaching criterion faster in session 2 than in session 1. Many participants were at ceiling in session 1 ($n = 32$), resulting in a violation of normality of differences. As such, median and interquartile range values are reported with a Wilcoxon test. In contrast to my previous experiments, I found that there was no significant difference between the number of trials to reach criterion in session 1 ($Median = 22$, $IQR = 3.5$) and session 2 ($Median = 22$, $IQR = 0$), V = 183.5, $p = .346$, $r = 0.17$. The lack of a difference here is likely a consequence of the new experimental design, in which the schemas were constructed of non-overlapping categories. This is likely to have made the task easier, resulting in a large number of participants at ceiling in session 1.

3.2.2.2 Encoding

To be included in the encoding analyses participants were required to be significantly above chance (50%) on the animacy judgement task. As such, the sample size for this analysis was $n = 46$. For the encoding analysis, the data were filtered to only include items that were used as the Cue and Cue+1 at test.

Mean accuracy and response times across the Congruency and Test manipulations can be seen in Table 3.1. A full summary of the mixed-effects model at encoding in Experiment 1 can been seen Table 3.2. The results of the encoding analysis are plotted in Figure 3.2.

Table 3.1: Mean accuracy and response times at encoding between Cue and Cue+1 across the three Congruency conditions in Experiment 3. Standard deviations are shown in parentheses.

Metric	Congruency	Cue	Cue+1
Accuracy (%)	Unrelated	75.04 (15.83)	82.17 (12.74)
	Congruent	85.65 (14.07)	90.09 (9.46)
	Incongruent	83.78 (13.69)	78.61 (9.97)
Response Time (ms)	Unrelated	702.03 (60.48)	675.95 (74.7)
	Congruent	630.57 (84.64)	618.06 (79.77)
	Incongruent	632.45 (77.83)	678.87 (69.59)

3.2.2.2.1 Accuracy

When investigating whether accuracy differed as a function of Congruency on Cue trials (where the categories were still congruent on both the Congruent and Incongruent trials) I found that accuracy was significantly greater in the Congruent ($\beta = 0.78$, *SE* = 0.08, z = 9.58, p <.001; prereg H1) and Incongruent ($\beta = 0.62$, *SE* = 0.08, z = 7.81, p <.001) conditions on Cue trials as compared to the Unrelated Cue trials. This suggests that participants were able to use their prior knowledge of sequence structure to facilitate accurate responses.

I then assessed differences between the Congruency conditions at Cue+1 (after the change had occurred in the Incongruent condition). Accuracy was greater in the Congruent condition as compared to the Unrelated ($\beta = 0.72$, *SE* = 0.09, z = 7.91, p <.001; prereg H1) and Incongruent conditions ($\beta = 0.96$, *SE* = 0.09, z = 10.84, p <.001). Performance in the Incongruent condition at Cue+1 was significantly lower than performance in the Unrelated condition ($\beta = -0.24$, *SE* = 0.08, z = -3.16, p = .002). This suggests that the Incongruent trial at Cue+1 had a negative impact on performance.

In line with the above findings, the main model output showed that there was a significant Congruency by Test interaction for the Incongruent vs. Unrelated comparison ($\beta = -0.86$, *SE* = 0.11, z = -7.81, p <.001), however, this interaction was not present for the Congruent vs. Unrelated comparison ($\beta = -0.06$, *SE* = 0.12, z = -0.53, p = .599).

When investigating differences between Cue and Cue+1 in each Congruency condition, I found that there was a significant increase in accuracy between Cue and Cue+1 in the Congruent ($\beta = 0.32$, *SE* = 0.12, z = 2.71, p = .007) and Unrelated conditions ($\beta = 0.38$, *SE* = 0.1, z = 3.77, p <.001). However, there was a significant decrease in accuracy from Cue to Cue+1 in the Incongruent condition ($\beta = -0.48$, *SE* = 0.1, z = -4.6, p <.001; prereg H2). Again, this indicates that the violation of the schemas at Cue+1 disrupted performance in the Incongruent condition, potentially due to a disruption of ongoing predictions of upcoming events.

3.2.2.2.2 Response time

During Cue trials, response times were significantly faster in the Congruent ($\beta = -72.83$, *SE* = 6.98, z = -10.44, p <.001; prereg H1) and Incongruent ($\beta = -71.73$, *SE* = 6.16, z = -11.65, p <.001) conditions as compared to Unrelated trials. This indicates that prior to the schema violation, prior knowledge was facilitating performance in the Congruent and Incongruent conditions.

At Cue+1 (after the change had occurred in the Incongruent condition), response times were faster in the Congruent condition as compared to both the Unrelated ($\beta = -56.69$, *SE* = 6.86, z = -8.27, p <.001; prereg H1) and Incongruent conditions ($\beta = -60.75$, *SE* = 5.5, z = -11.04, p <.001). However, there was no difference in response times between the Incongruent and Unrelated trials at Cue+1 ($\beta = 4.06$, *SE* = 6.12, z = 0.66, p = .507).

Once again there was a significant Congruency by Test interaction when comparing the Incongruent and Unrelated conditions ($\beta = 75.79$, *SE* = 6.76, z = 11.21, p <.001). However, in contrast to previous experiments, there was now also an interaction when comparing the Congruent and Unrelated condition ($\beta = 16.14$, *SE* = 6.64, z = 2.43, p = .015).

When comparing response times between Cue and Cue+1 in each Congruency condition there was no difference in response times in the Congruent condition ($\beta = -11.77$, *SE* = 6.85, z = -1.72, p = .086). However, response times were faster at Cue+1 relative to Cue in the Unrelated condition ($\beta = -27.91$, *SE* = 7.03, z = -3.97, p <.001). This indicates that the Congruent vs. Unrelated interaction with Test was driven by differences between Cue and Cue+1 in the Unrelated condition with faster responses at Cue+1. Most notably, response times were slower at Cue+1 in the Incongruent condition ($\beta = 47.88$, *SE* = 6.97, z = 6.87, p <.001; prereg H2). Together, these results suggest that ongoing predictions based on schemas were indeed disrupted by the violation at Cue+1.

Overall, the effects observed in the encoding analysis of Experiment 3 were, with minor variations, highly consistent with the pattern of results observed in Experiments 1 and 2. This suggests that these effects are highly robust and replicable across experiments.

Table 3.2: Summary of the mixed-effects model at encoding for both accuracy and response time in Experiment 3. P-values are uncorrected and therefore may differ from those in the text.

Metric	Fixed effect	Estimate	SE	z/t	p
Accuracy (%)	Intercept	1.29	0.15	8.81	<.001
	Congruent (vs. unrelated)	0.78	0.08	9.58	<.001
	Incongruent (vs. unrelated)	0.62	0.08	7.81	<.001
	Test (Cue vs. Cue+1)	0.38	0.10	3.77	<.001
	Congruent x Test	-0.06	0.12	-0.53	.599
	Incongruent x Test	-0.86	0.11	-7.81	<.001
Response Time (ms)	Intercept	702.69	9.73	72.19	<.001
	Congruent (vs. unrelated)	-72.83	6.98	-10.44	<.001
	Incongruent (vs. unrelated)	-71.73	6.16	-11.65	<.001
	Test (Cue vs. Cue+1)	-27.91	7.03	-3.97	<.001
	Congruent x Test	16.14	6.64	2.43	.015
	Incongruent x Test	75.79	6.76	11.21	<.001

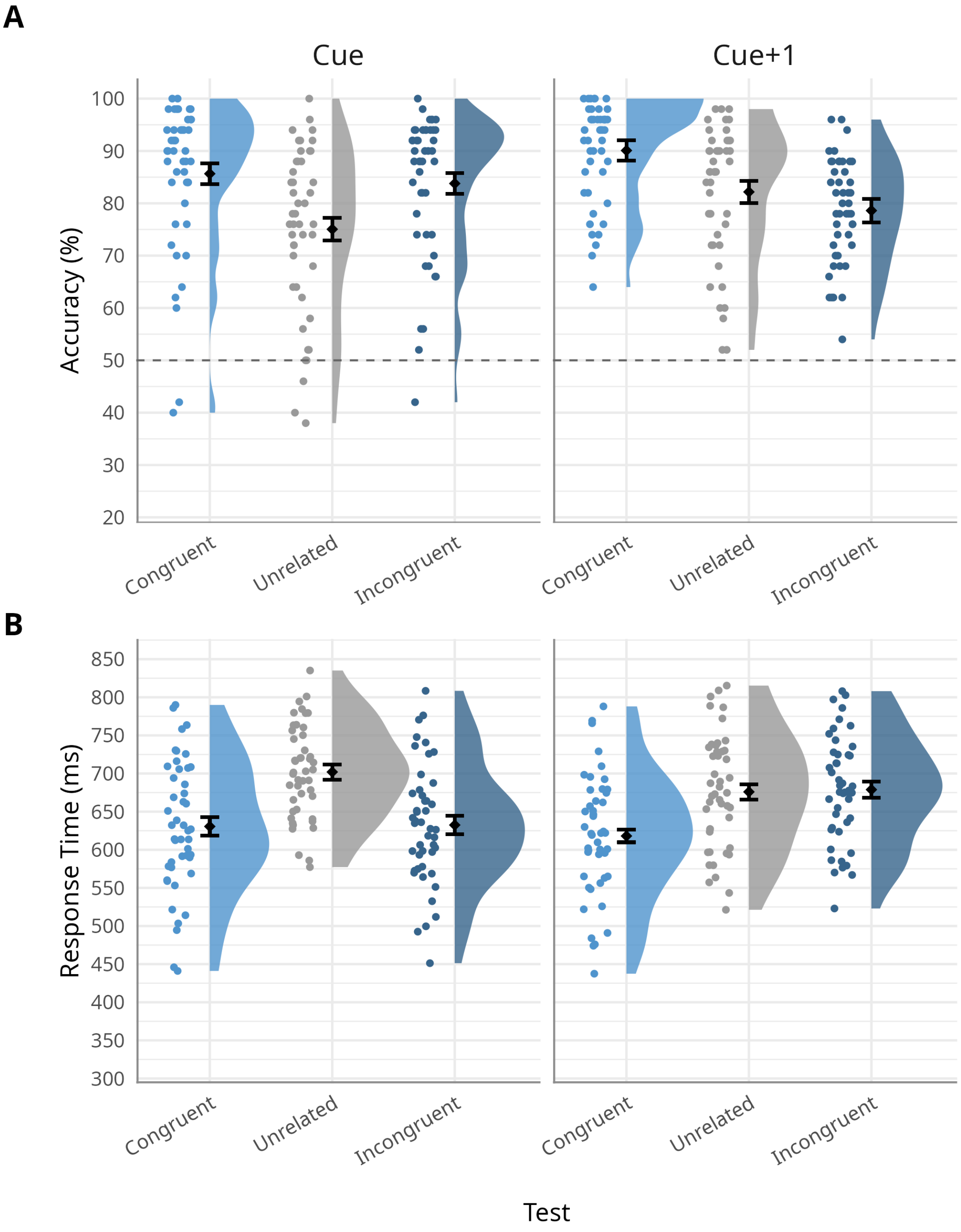

Figure 3.2: (A) Mean accuracy (%) and (B) response time (ms) across the Congruency and Test manipulations at encoding in Experiment 3. Diamonds show sample means. Density and scatter plots show the distribution of participant means. Error bars represent 95% confidence intervals.

3.2.2.3 Retrieval

Mean accuracy across the Congruency and Test manipulations at retrieval in Experiment 3 can be seen in Table 3.3. A full summary of the mixed-effects model at retrieval can be seen in Table 3.4. The results of this analysis are plotted in Figure 3.3.

Table 3.3: Mean percentage accuracy at retrieval between Cue+1 and Cue+2 across the three Congruency conditions at the category, exemplar, and conditional exemplar level in Experiment 3. Standard deviations are shown in parentheses.

Level	Congruency	Cue+1	Cue+2
Category	Unrelated	59.13 (11.1)	58.7 (10.61)
	Congruent	80.78 (11.43)	79.78 (10.78)
	Incongruent	55.52 (13.49)	59.57 (11.78)
Exemplar	Unrelated	42.74 (14.31)	42.22 (15.13)
	Congruent	56.57 (17.05)	54.52 (16.18)
	Incongruent	40.52 (16.36)	39.7 (15.82)
Conditional Exemplar	Unrelated	74.9 (10.72)	74.36 (13.22)
	Congruent	73.57 (11.76)	70.96 (12.24)
	Incongruent	73.73 (12.65)	67.76 (14.54)

3.2.2.3.1 Category level

To be included in the category-level retrieval analysis participants were required to be significantly above chance (50%) in terms of category-level accuracy collapsed across Congruency and Test. The final sample for the category analysis was $n = 46$.

Collapsing across Test, I found that accuracy at the category level was greater in the Congruent condition relative to the Unrelated ($\beta = 1.16$, $SE = 0.1$, $z = 11.25$, $p < .001$) and Incongruent conditions ($\beta = 1.21$, $SE = 0.09$, $z = 12.72$, $p < .001$), however, there was no difference between the Incongruent and Unrelated condition ($\beta = -0.05$, *SE* = 0.06, *z* = -0.78, *p* = .437).

There was also a significant Congruency by Test interaction for the Incongruent vs. Unrelated contrast ($\beta = 0.19$, *SE* = 0.09, *z* = 2.23, *p* = .026). When breaking the model down by Test, at Cue+1, category-level accuracy was greater on Congruent trials than Unrelated ($\beta = 1.18$, *SE* = 0.11, *z* = 10.36, *p* <.001; prereg H3) and Incongruent trials ($\beta = 1.33$, *SE* = 0.11, *z* = 12.42, *p* <.001). However, there was no difference between the Incongruent and Unrelated conditions ($\beta = -0.14$, *SE* = 0.07, *z* = -1.93, *p* = .053; prereg H3). At Cue+2, I saw the same pattern of results as at Cue+1. Accuracy was higher in the Congruent condition relative to both the Unrelated ($\beta = 1.14$, *SE* = 0.11, *z* = 9.98, *p* <.001) and Incongruent conditions ($\beta = 1.09$, *SE* = 0.11, *z* = 10.19, *p* <.001). However, there was no difference between the Incongruent and Unrelated conditions ($\beta = 0.05$, *SE* = 0.07, *z* = 0.66, *p* = .509). This indicates that the interaction was driven by numerical differences between the Incongruent and Unrelated conditions, the Incongruent condition was numerically lower at Cue+1 but higher at Cue+2, however, neither of these effects reached statistical significance when analysed on their own.

Overall, this pattern of results indicates that the use of non-overlapping schemas did not produce an incongruency advantage at the category level. However, in line with previous experiments, the congruency advantage was present, indicating its high replicability.

3.2.2.3.2 Exemplar level

To be included in the exemplar-level retrieval analysis participants were required to be significantly above chance (25%) in terms of exemplar-level accuracy collapsed across Congruency and Test manipulations. This left a sample size of $n = 46$.

Overall, the effects at the exemplar level mirrored those seen at the category level. Collapsed across Test, accuracy was greater in the Congruent condition as compared to the

Unrelated ($\beta = 0.57$, *SE* = 0.06, $z = 9.01$, p <.001) and Incongruent conditions ($\beta = 0.68$, *SE* = 0.07, $z = 9.28$, p <.001). But there was no difference between the Incongruent and Unrelated conditions ($\beta = -0.11$, *SE* = 0.06, z = -1.81, p = .070). In contrast to the category-level model, there were no interactions between Congruency and Test at the exemplar level.

When breaking the exemplar model down by Test, at Cue+1, accuracy was higher in the Congruent condition relative to the Unrelated ($\beta = 0.61$, *SE* = 0.08, $z = 7.85$, p <.001) and Incongruent conditions ($\beta = 0.71$, *SE* = 0.09, $z = 8.25$, p <.001). However, there was no difference in accuracy between the Incongruent and Unrelated conditions ($\beta = -0.1$, *SE* = 0.07, z = -1.35, p = .176). I observed the same pattern of results at Cue+2. Accuracy was higher in the Congruent condition relative to the Unrelated ($\beta = 0.54$, *SE* = 0.08, z = 6.97, p <.001) and Incongruent conditions ($\beta = 0.65$, *SE* = 0.09, $z = 7.63$, p <.001), but there was no difference between the Incongruent and Unrelated conditions ($\beta = -0.11$, *SE* = 0.07, z = -1.55, p = .121).

3.2.2.3.3 Conditional exemplar level

To be included in the conditional exemplar-level analysis participants were required to be above chance performance at the exemplar level when only considering trials on which category-level information was accurately recalled (50%). This resulted in a final sample size of $n = 38$.

There were no significant effects when collapsing across Test. However, there was a significant Congruency by Test interaction for the Incongruent vs Unrelated contrast ($\beta = -0.29$, *SE* = 0.14, z = -2.08, p = .037). When breaking the model down by Test, the only significant finding was that accuracy was lower in the Incongruent condition relative to the Unrelated baseline at Cue+2 ($\beta = -0.34$, *SE* = 0.1, z = -3.54, p = .001), indicating that the interaction was driven by a reduction in performance in the Incongruent condition at Cue+2.

Overall, this indicates that constructing the schemas from non-overlapping categories did not have any impact on the emergence of exemplar-level schema effects.

Table 3.4: Summary of the mixed-effects model at retrieval for the category, exemplar, and conditional exemplar-level analyses in Experiment 3. P-values are uncorrected and therefore may differ from those in the text.

Level	Fixed effect	Estimate	SE	z	p
Category	Intercept	0.38	0.07	5.37	<.001
	Congruent (vs. unrelated)	1.18	0.11	10.36	<.001
	Incongruent (vs. unrelated)	-0.14	0.07	-1.93	.053
	Test (Cue+1 vs. Cue+2)	-0.02	0.06	-0.30	.760
	Congruent x Test	-0.05	0.10	-0.50	.618
	Incongruent x Test	0.19	0.09	2.23	.026
Exemplar	Intercept	-0.31	0.09	-3.23	.001
	Congruent (vs. unrelated)	0.61	0.08	7.85	<.001
	Incongruent (vs. unrelated)	-0.10	0.07	-1.35	.176
	Test (Cue+1 vs. Cue+2)	-0.02	0.06	-0.37	.710
	Congruent x Test	-0.07	0.09	-0.78	.438
	Incongruent x Test	-0.01	0.09	-0.17	.867
Conditional Exemplar	Intercept	1.16	0.11	10.23	<.001
	Congruent (vs. unrelated)	0.00	0.10	0.01	.992
	Incongruent (vs. unrelated)	-0.06	0.09	-0.70	.484
	Test (Cue+1 vs. Cue+2)	-0.05	0.10	-0.52	.606
	Congruent x Test	-0.12	0.13	-0.89	.371
	Incongruent x Test	-0.29	0.14	-2.08	.037

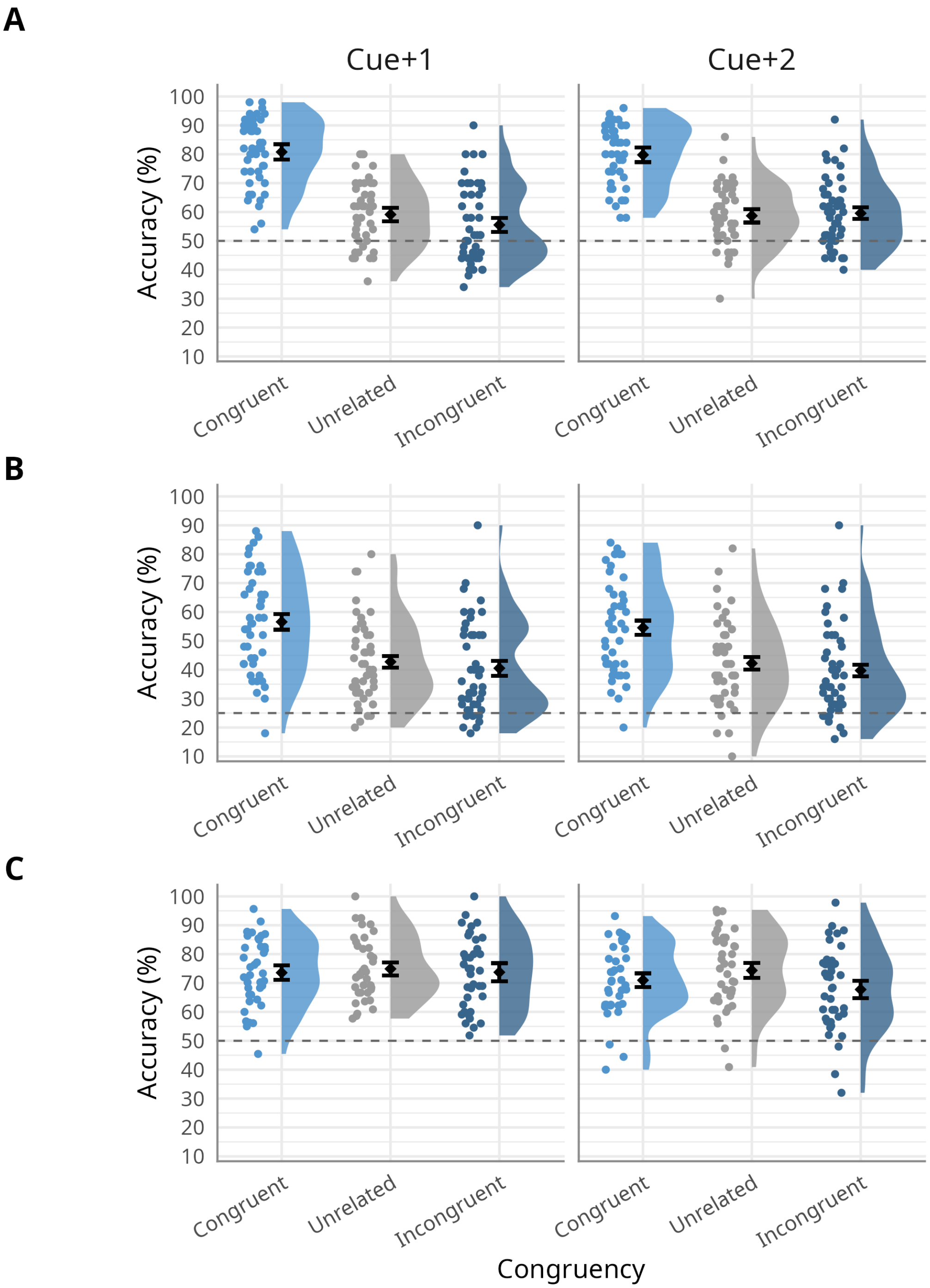

Figure 3.3: Mean accuracy (%) at retrieval across the Congruency and Test manipulations at (A) category level, (B) exemplar level, and (C) conditional exemplar level in Experiment 3. Diamonds show the sample means. Density and scatter plots show the distribution of participant means. Error bars represent 95% within-subject confidence intervals. 116

3.2.2.4 Exploratory analysis

In Experiments 2 and 3, I observed no incongruency effect. However, I did observe evidence of prediction errors at encoding. Moreover, the exploratory analyses of Experiment 2 indicated that these prediction errors were primarily occurring at the animacy level, with no prediction errors being observed when a violation occurred at the category level only. This raises the question as to whether the presence or absence of a violation at the animacy level could impact the emergence of an incongruency effect at retrieval. That is to say, is an incongruency effect at retrieval dependent on a violation at the animacy level? To address this hypothesis, I broke the data down by whether the violation at Cue+1 violated the schema at the animacy level (and the category level) or only violated the schema at the category level. I then investigated whether the emergence of an incongruency advantage at the category level at Cue+1 during retrieval was dependent on the presence of absence of an animacy violation at encoding. I conducted this analysis as 2x3 mixed ANOVA with the combined data from Experiments 2 and 3. The overall design of this model was a within-subjects predictor, Violation, with three levels: Unrelated (baseline), Category-only violation, Animacy violation, and a between-subjects predictor, Experiment, with two levels: Experiment 2 and Experiment 3.

This model revealed a main effect of Experiment, $F(1,\ 93) = 6.4$, $p = .013$, $\eta^2 = 0.04$, but no main effect of Violation, $F(1.81,\ 168.33) = 1.39$, $p = .252$, $\eta^2 < .01$, and no interaction, $F(1.81,\ 168.33) = 0.64$, $p = .512$, $\eta^2 < .01$. Together, these results suggest that whether a violation occurred at the animacy level had no effect on whether there was a difference between the Unrelated and Incongruent conditions for category-level performance at Cue+1.

3.2.3 Discussion

In Experiment 3, I successfully replicated the major findings from Experiment 2. When using four-item sequences, schema congruency effects were observed at the category level. Moreover, these congruency effects were preceded by evidence of predictions at encoding. Performance on the animacy judgements was higher in the Congruent and Incongruent conditions relative to the Unrelated condition on Cue trials. However, at Cue+1, performance was reduced in the Incongruent condition at Cue+1, suggesting a disruption of ongoing expectations. Crucially, I also replicated the finding that schema incongruency had no effect on memory retention. At both the category and exemplar level, there was no difference in recall between the Incongruent condition and the Unrelated baseline.

This latter finding suggests that the lack of an incongruency effect in the preceding chapter was not due to interference caused by overlapping schema structures. Overall, this leads to the possibility that temporal reordering is not effective at producing incongruency effects in the present paradigm. Importantly, this finding has implications for the nature of the schemas that are being generated in this paradigm. As previously discussed, given the spatiotemporal nature of episodic memories, we would expect some schemas to be sensitive to chronology. However, the results here suggest that, even when temporal order information is highly relevant, this may not always be the case. I will return to this point in more detail in the general discussion. For now, the crucial point is that the schemas being generated may not contain any information regarding temporal order. Instead, the schemas may simply contain information regarding content, e.g., what categories are included in the sequence.

Notably, one reason to doubt this interpretation is that it would predict a “congruency” advantage in the Incongruent condition. That is, if all categories are equally predicted, then any presentation order is technically schema-congruent. Given that no memory advantage was seen in the Incongruent condition, this problem requires an explanation. One possibility is that the congruency effects in the Congruent condition are purely driven by retrieval-based strategies. If this were the case, there would be no reason to expect the predictions occurring in the Incongruent condition to produce a congruency advantage, as they would need to exert their influence at encoding. A second possibility is that temporal order information is indeed represented in some way. Rather than each category being predicted equally, multiple categories may be predicted, but their degree of activation may be relative to their proximity, with more proximal categories being predicted more strongly than distal ones. If this were the case, temporal reordering would still be schema-incongruent, preventing any congruency advantages in the Incongruent condition. However, the partial activation of categories multiple steps into the future may be enough to suppress a prediction error strong enough to drive new learning when those categories occur earlier than expected.

Experiment 4 aimed to address this hypothesis by using an incongruency manipulation that manipulates content, rather than temporal order. By introducing a category that is completely unexpected in a given schema, not even partial activation of that category should be possible. If an incongruency advantage could be demonstrated in such a case, this would provide evidence that predictions are indeed occurring multiple steps into the future with temporal order represented as stronger activation for more proximal items. Experiment 4 used the same design as Experiment 3, however, incongruency was now defined as the introduction of a category that was completely unexpected for a given context.

3.3 Behavioural Experiment 4

3.3.1 Methods

3.3.1.1 Participants

Fifty-two participants were recruited via SONA systems. Three participants were excluded due to a software error leading to an incomplete dataset. This left a final sample of $N = 49$ participants (40 females, 8 males, 1 prefer not to say; $M \pm SD$ age in years $= 20.62 \pm 2.92$). All participants were fluent English speakers with normal or corrected-to-normal vision. Participants gave informed consent for the experiment, which was approved by the Research Ethics Committee of the Department of Psychology, University of York. Participants were reimbursed for their time with either course credit or a £15 Amazon voucher.

3.3.1.2 Procedure

Experiment 4 used the same design as Experiment 3. However, incongruency was now defined as an introduction of a category that is completely unexpected for a context (rather than the reordering of expected categories). I achieved this by introducing a category from the other schema during Incongruent trials. This controlled for overall category familiarity while violating schematic expectations at the level of the category. For example, if the Congruent schemas were: (red) *face-building-dog-tool*, (blue) *furniture-car-plant-insect*, then an Incongruent sequence could be: (red) *face-building-car-insect*, (blue) *furniture-face-dog-insect*. A schematic outlining this new manipulation can be seen in Figure 3.4.

3.3.1.3 Analysis approach

The analysis approach used in Experiment 4 mirrored that used in Experiment 3 and was pre-registered on the Open Science Framework (Exp 4 pre-registration).

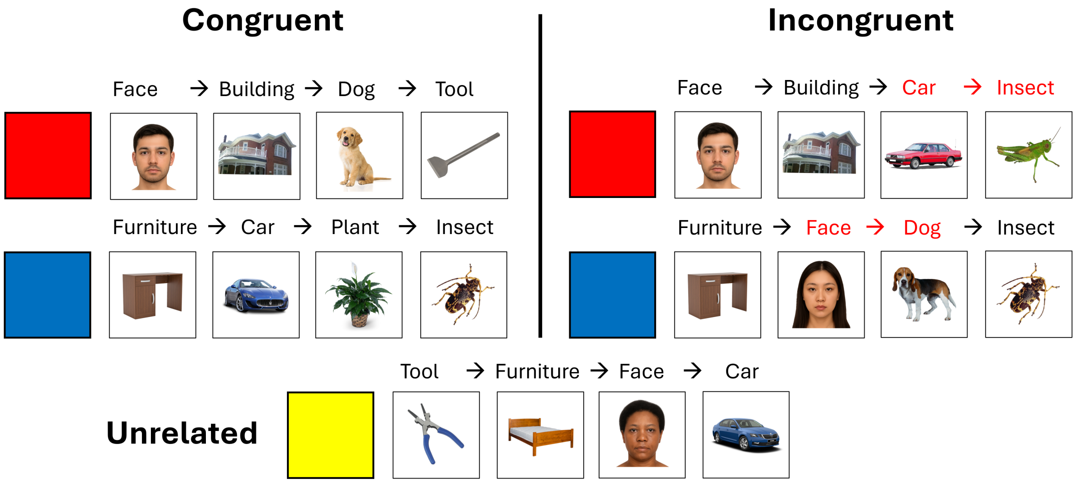

Figure 3.4: Schematic depicting the incongruency manipulation in Experiment 4. The Congruent and Unrelated conditions were identical to those used in Experiment 3. In the Incongruent condition, incongruency was defined as the introduction of a category that was completely unexpected in that context. For example, in the red schema, an Incongruent trial would be the introduction of any of the categories from the blue schema. If incongruency occurred on the second trial, as in the blue example above, the final item in the sequence would still be schema-congruent.

3.3.2 Results

3.3.2.1 Pre-learning

First, I ran an analysis on trials-to-criterion in each session to assess whether participants were retaining knowledge from the first pre-learning session. As with Experiment 3, a large number of participants were at ceiling at session 1 ($n = 31$), resulting in a violation of normality of differences. As such, median and interquartile range values with a Wilcoxon test are reported. Although the median values were identical in session 1 (*Median* = 22, *IQR* = 7) and session 2 (*Median* = 22, *IQR* = 0), the Wilcoxon test reached statistical significance V = 173.5, $p = .011$, $r = 0.43$, indicating that variability differed between the sessions. This indicates that although many participants showed no difference between sessions, there was a group level effect by which more trials were required to reach criterion on the first session. This suggests that participants were retaining knowledge from the first pre-learning session.

3.3.2.2 Encoding

To be included in the encoding analyses participants were required to be significantly above chance (50%) on the animacy judgement task. Therefore, the sample size for this analysis was $n = 49$. For the encoding analysis, the data were filtered to only include items that were used as the Cue and Cue+1 at test.

Mean accuracy and response times across the Congruency and Test manipulations can be seen in Table 3.5. A full summary of the mixed-effects model at encoding in Experiment 1 can been seen Table 3.6. The results of the encoding analysis are plotted in Figure 3.5.

Table 3.5: Mean accuracy and response times at encoding between Cue and Cue+1 across the three Congruency conditions in Experiment 4. Standard deviations are shown in parentheses.

Metric	Congruency	Cue	Cue+1
Accuracy (%)	Unrelated	76.86 (14.51)	81.18 (13.19)
	Congruent	85.47 (12.95)	89.02 (9.13)
	Incongruent	84.57 (12.21)	74.86 (14.07)
Response Time (ms)	Unrelated	701.91 (64.89)	680.36 (66.1)
	Congruent	629.93 (75.06)	616.55 (81.26)
	Incongruent	640.46 (77.54)	712.54 (60.56)

3.3.2.2.1 Accuracy

First, I investigated whether accuracy differed as a function of Congruency on Cue trials (where the categories were still congruent on both the congruent and incongruent trials). Accuracy was significantly greater on Congruent ($\beta = 0.64$, $SE = 0.08$, $z = 8.12$, $p < .001$; prereg H1) and Incongruent ($\beta = 0.56$, $SE = 0.08$, $z = 7.23$, $p < .001$) Cue trials as compared to the Unrelated Cue trials. This suggests that prior knowledge was facilitating performance in the Congruent and Incongruent conditions.

On Cue+1 trials accuracy was significantly greater in the Congruent condition as compared to both the Unrelated ($\beta = 0.67$, $SE = 0.09$, $z = 7.89$, $p < .001$; prereg H1) and Incongruent conditions ($\beta = 1.07$, $SE = 0.08$, $z = 13.05$, $p < .001$). Crucially, accuracy was now lower in the Incongruent condition as compared to the Unrelated condition ($\beta = -0.4$, $SE = 0.07$, $z = -5.57$, $p < .001$). These results indicate the violation of the schemas at Cue+1 negatively impacted performance in the Incongruent condition.

In line with the interpretation that performance varied in the incongruent condition across the Test manipulation, there was a significant Congruency by Test interaction when comparing the Incongruent and Unrelated conditions ($\beta = -0.97$, *SE* = 0.11, z = -9.09, p <.001). Moreover, there was no Congruency by Test interaction for the Congruent vs. Unrelated comparison ($\beta = 0.03$, *SE* = 0.12, z = 0.27, p = .788).

When comparing Cue vs. Cue+1 in each Congruency condition I found a significant difference in all three comparisons. Performance significantly improved between Cue and Cue+1 trials in both the Congruent ($\beta = 0.25$, *SE* = 0.1, z = 2.38, p = .017) and Unrelated condition ($\beta = 0.22$, *SE* = 0.09, z = 2.41, p = .016). However, there was a significant decrease in performance between Cue and Cue+1 in the Incongruent condition ($\beta = -0.75$, *SE* = 0.09, z = -8.18, p <.001; prereg H2). The decrease in performance in the Incongruent condition supports the notion that participants' ongoing predictions of upcoming events were perturbed by the sequence violations.

3.3.2.2.2 Response time

During Cue trials, response times were significantly faster in the Congruent ($\beta = -72.44$, *SE* = 7.42, z = -9.76, p <.001; prereg H1) and Incongruent ($\beta = -62.09$, *SE* = 5.08, z = -12.22, p <.001) conditions as compared to Unrelated trials indicating that schema knowledge was facilitating performance.

After the schema violation at Cue+1, response times were faster in the Congruent condition as compared to both the Unrelated ($\beta = -65.08$, *SE* = 7.35, z = -8.85, p <.001; prereg H1) and Incongruent conditions ($\beta = -95.55$, *SE* = 6.44, z = -14.83, p <.001). However, the response times in the Incongruent condition at Cue+1 were slower than response times in the Unrelated condition ($\beta = 30.47$, *SE* = 5.15, z = 5.91, p <.001) indicating that the schema violation perturbed performance.

Once again there was a significant Congruency by Test interaction when comparing the Incongruent and Unrelated conditions ($\beta = 92.56$, *SE* = 6.49, z = 14.25, p <.001) but no interaction when comparing the Congruent and Unrelated condition ($\beta = 7.36$, $SE = 6.35$, $z = 1.16$, $p = .246$).

When comparing response times between Cue and Cue+1 in each Congruency condition I observed a significant increase in response speed in both the Congruent ($\beta = -12.8$, $SE = 4.38$, $z =$ -2.93, $p = .003$) and Unrelated conditions ($\beta = -20.16$, $SE = 4.6$, $z =$ -4.38, $p <.001$). However, in the Incongruent condition, response times were significantly slower at Cue+1 ($\beta = 72.4$, $SE = 4.58$, $z = 15.79$, $p <.001$; prereg H2).

Table 3.6: Summary of the mixed-effects model at encoding for both accuracy and response time in Experiment 4. P-values are uncorrected and therefore may differ from those in the text.

Metric	Fixed effect	Estimate	SE	z/t	p
Accuracy (%)	Intercept	1.39	0.14	10.24	<.001
	Congruent (vs. unrelated)	0.64	0.08	8.12	<.001
	Incongruent (vs. unrelated)	0.56	0.08	7.23	<.001
	Test (Cue vs. Cue+1)	0.22	0.09	2.41	.016
	Congruent x Test	0.03	0.12	0.27	.788
	Incongruent x Test	-0.97	0.11	-9.09	<.001
Response Time (ms)	Intercept	701.36	9.30	75.44	<.001
	Congruent (vs. unrelated)	-72.44	7.42	-9.76	<.001
	Incongruent (vs. unrelated)	-62.09	5.08	-12.22	<.001
	Test (Cue vs. Cue+1)	-20.16	4.60	-4.38	<.001
	Congruent x Test	7.36	6.35	1.16	.246
	Incongruent x Test	92.56	6.49	14.25	<.001

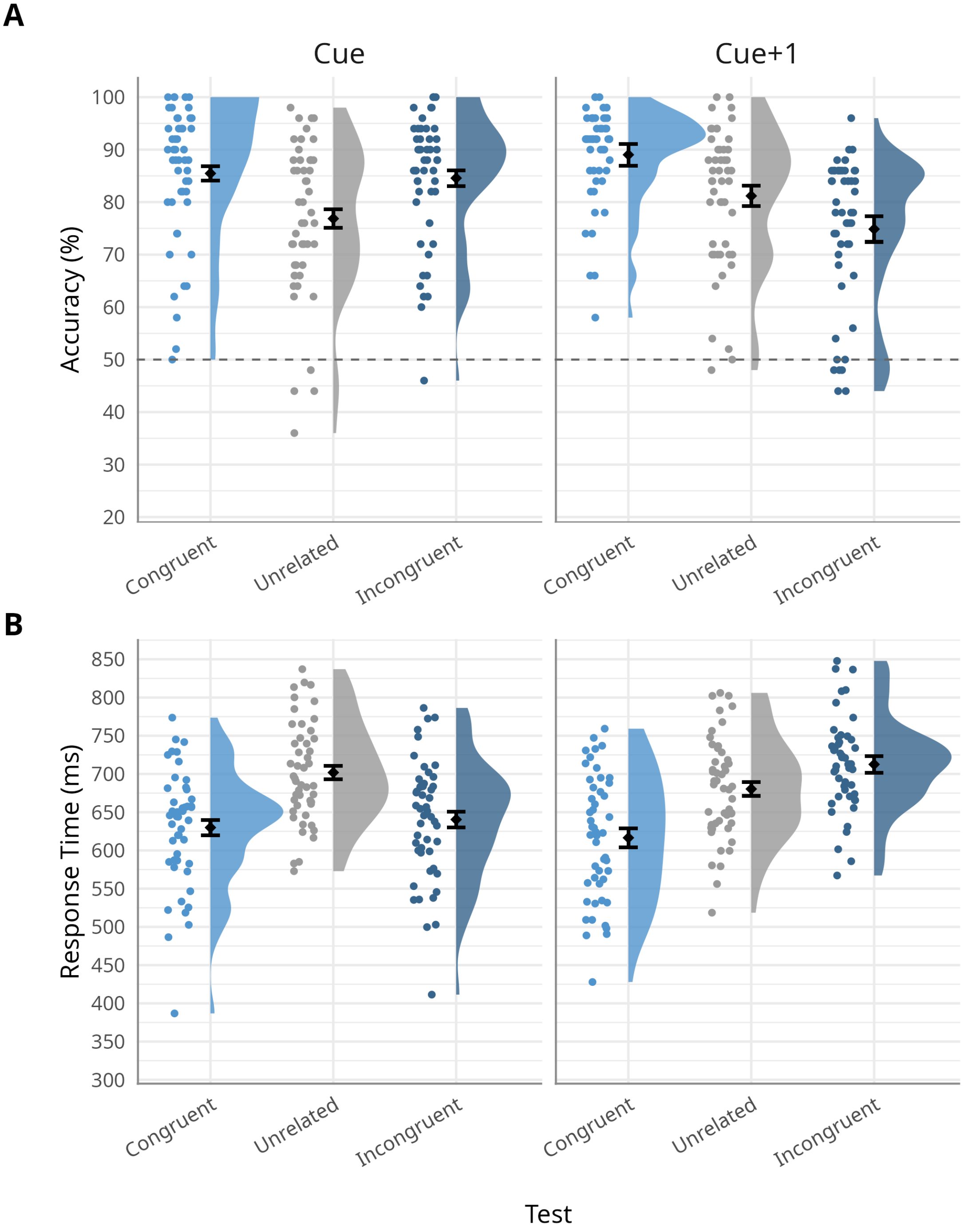

Figure 3.5: (A) Mean accuracy (%) and (B) response time (ms) across the Congruency and Test manipulations at encoding in Experiment 4. Diamonds show sample means. Density and scatter plots show the distribution of participant means. Error bars represent 95% confidence intervals.

3.3.2.3 Retrieval

Mean accuracy across the Congruency and Test manipulations at retrieval in Experiment 4 can be seen in Table 3.7. A full summary of the mixed-effects model at retrieval can be seen in Table 3.8. The results of this analysis are plotted in Figure 3.6.

Table 3.7: Mean percentage accuracy at retrieval between Cue+1 and Cue+2 across the three Congruency conditions at the category, exemplar, and conditional exemplar level in Experiment 4. Standard deviations are shown in parentheses.

Level	Congruency	Cue+1	Cue+2
Category	Unrelated	60.4 (11.71)	61.47 (9.88)
	Congruent	82.09 (10.45)	81.16 (10.14)
	Incongruent	68.31 (11.31)	63.64 (10.93)
Exemplar	Unrelated	43.28 (16.7)	44.13 (14.96)
	Congruent	59.53 (16.68)	57.66 (16.66)
	Incongruent	48.13 (17.8)	43.96 (15.83)
Conditional Exemplar	Unrelated	74.57 (12.21)	74.79 (12.17)
	Congruent	75.48 (12.11)	74.29 (11.51)
	Incongruent	72.17 (15.01)	71.55 (13.55)

3.3.2.3.1 Category level

To be included in the category-level retrieval analysis participants were required to be significantly above chance (50%) in terms of category-level accuracy collapsed across Congruency and Test. This resulted in a final sample size of $n = 45$

Collapsing across Test, I saw that accuracy at the category level was greater in the Congruent condition relative to the Unrelated ($\beta = 1.08$, $SE = 0.05$, $z = 21.66$, $p <.001$) and Incongruent conditions ($\beta = 0.85$, $SE = 0.05$, $z = 16.9$, $p <.001$). Crucially, and in contrast to the previous experiments, there was now a significant difference between the Incongruent and Unrelated condition ($\beta = 0.23$, $SE = 0.04$, $z = 5.1$, $p < .001$) with accuracy being higher in the Incongruent condition.

I also observed a significant Incongruency by Test interaction ($\beta = -0.26$, $SE = 0.09$, $z = -2.94$, $p = .003$). When breaking the model down by Test, at Cue+1, category-level accuracy was greater on Congruent trials than Unrelated trials ($\beta = 1.14$, $SE = 0.07$, $z = 16.04$, $p < .001$; prereg H3) and Incongruent trials ($\beta = 0.78$, $SE = 0.07$, $z = 10.76$, $p < .001$). However, category-level accuracy was higher in the Incongruent condition relative to the Unrelated condition ($\beta = 0.36$, $SE = 0.06$, $z = 5.65$, $p < .001$); prereg H3). At Cue+2, performance was still greater in the Congruent condition relative to the Unrelated ($\beta = 1.03$, $SE = 0.07$, $z = 14.62$, $p < .001$) and Incongruent conditions ($\beta = 0.93$, $SE = 0.07$, $z = 13.19$, $p < .001$). However, there was no difference between the Incongruent and Unrelated condition at Cue+2 ($\beta = 0.1$, $SE = 0.06$, $z = 1.54$, $p = .123$).

Overall, these results replicated the Congruency advantage observed in the previous experiments, further emphasising the robustness of the congruency effect. Crucially, I also observed a significant incongruency advantage indicating that the critical U-shaped function of schema congruency was present in Experiment 4. Notably, however, the incongruency advantage is only present at Cue+1 and at the category level. The presence of an incongruency advantage for categorical violations has important implications for the nature of the predictions that are taking place which I will return to in the discussion section.

3.3.2.3.2 Exemplar level

To be included in the exemplar-level retrieval analysis participants were required to be significantly above chance (25%) in terms of exemplar-level accuracy collapsed across Congruency and Test manipulations. This resulted in a final sample size of $n = 47$.

Collapsed across Test, accuracy was greater in the Congruent condition as compared to the Incongruent ($\beta = 0.56$, *SE* = 0.06, $z = 9.48$, p <.001) and Unrelated condition ($\beta = 0.66$, *SE* = 0.05, $z = 13.9$, p <.001). Likewise, exemplar-level accuracy was greater in the Incongruent condition as compared to the Unrelated condition ($\beta = 0.11$, *SE* = 0.05, $z = 2.14$, p = .032).

Once again, I observed a significant Incongruency by Test interaction ($\beta = -0.22$, *SE* = 0.09, z = -2.57, p = .010). When breaking the model down by Test, at Cue+1, exemplar-level accuracy was greater in the Congruent condition as compared to both the Unrelated ($\beta = 0.73$, *SE* = 0.06, $z = 11.18$, p <.001) and the Incongruent condition ($\beta = 0.51$, *SE* = 0.07, $z = 6.92$, p <.001). Likewise, performance was also greater in the Incongruent condition relative to the Unrelated condition ($\beta = 0.22$, *SE* = 0.07, $z = 3.31$, p <.001). At Cue+2, accuracy was greater in the Congruent condition relative to the Unrelated ($\beta = 0.6$, *SE* = 0.06, $z = 9.32$, p <.001) and Incongruent condition ($\beta = 0.61$, *SE* = 0.07, $z = 8.29$, p <.001), however, there was no difference between the Incongruent and Unrelated condition ($\beta <$ .01, *SE* = 0.07, z = -0.09, p = .930).

3.3.2.3.3 Conditional exemplar level

To be included in the conditional exemplar-level analysis participants were required to be above chance performance at the exemplar level when only considering trials on which category-level information was accurately recalled (50%). This resulted in a final sample size of $n = 39$.

When collapsing across Test, there was no difference between the Congruent and Unrelated condition ($\beta <$ .01, *SE* = 0.06, $z = 0.01$, p = .988). However, performance in the Congruent condition was higher than in the Incongruent condition ($\beta = 0.15$, *SE* = 0.06, $z = 2.49$, p = .039). Likewise, performance in the Incongruent condition was lower than performance in the Unrelated condition ($\beta = -0.15$, *SE* = 0.07, z = -2.28, p = .045). This suggests a potential incongruency disadvantage for conditional exemplar-level performance. However, when the model was broken down by Test to evaluate effects at Cue+1 and Cue+2 independently, there were no significant effects in any of the comparisons with $p > .16$ in all cases after correction for multiple comparisons. That is, the incongruency disadvantage was only significant when data was collapsed across Cue+1 and Cue+2.

Table 3.8: Summary of the mixed-effects model at retrieval for the category, exemplar, and conditional exemplar-level analyses in Experiment 4. P-values are uncorrected and therefore may differ from those in the text.

Level	Fixed effect	Estimate	SE	z	p
Category	Intercept	0.45	0.08	5.55	<.001
	Congruent (vs. unrelated)	1.14	0.07	16.04	<.001
	Incongruent (vs. unrelated)	0.36	0.06	5.65	<.001
	Test (Cue+1 vs. Cue+2)	0.05	0.06	0.75	.453
	Congruent x Test	-0.11	0.10	-1.11	.268
	Incongruent x Test	-0.26	0.09	-2.94	.003
Exemplar	Intercept	-0.28	0.10	-2.82	.005
	Congruent (vs. unrelated)	0.73	0.06	11.18	<.001
	Incongruent (vs. unrelated)	0.22	0.07	3.31	<.001
	Test (Cue+1 vs. Cue+2)	0.04	0.06	0.62	.538
	Congruent x Test	-0.12	0.09	-1.41	.159
	Incongruent x Test	-0.22	0.09	-2.57	.010
Conditional Exemplar	Intercept	1.18	0.12	9.60	<.001
	Congruent (vs. unrelated)	0.04	0.09	0.44	.662
	Incongruent (vs. unrelated)	-0.13	0.09	-1.34	.180
	Test (Cue+1 vs. Cue+2)	0.02	0.10	0.15	.878
	Congruent x Test	-0.08	0.13	-0.61	.544
	Incongruent x Test	-0.05	0.13	-0.40	.691

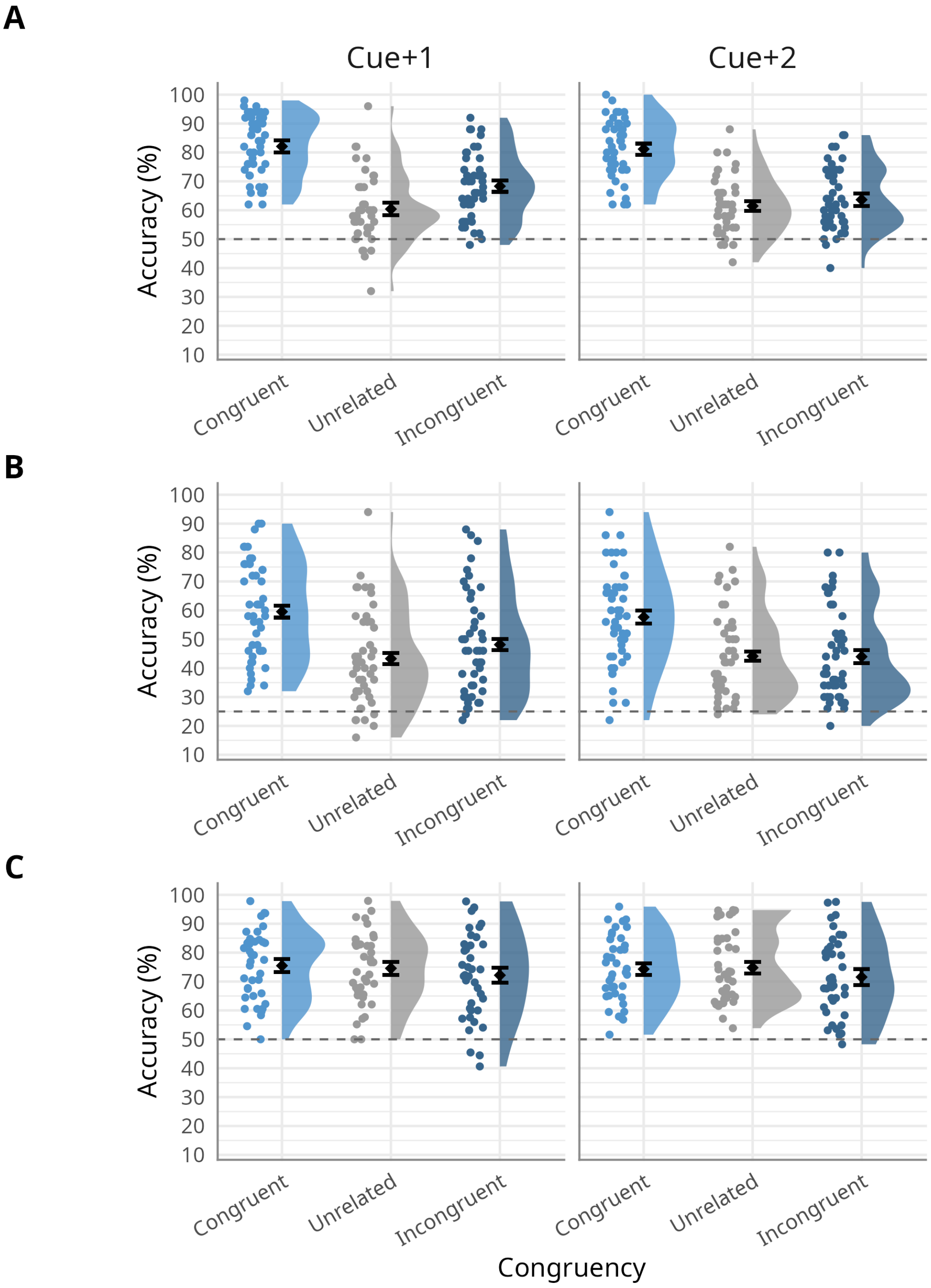

Figure 3.6: Mean accuracy (%) at retrieval across the Congruency and Test manipulations at (A) category level, (B) exemplar level, and (C) conditional exemplar level in Experiment 4. Diamonds show the sample means. Density and scatter plots show the distribution of participant means. Error bars represent 95% within-subject confidence intervals.

3.3.2.4 Exploratory analyses

3.3.2.4.1 Retrieval-based accounts of the incongruency advantage

In Experiment 4, I observed an incongruency advantage at the category level. This is important for several reasons, one of which being that it provides evidence that the effects observed at retrieval are a consequence of encoding-based processes, as would be predicted by neurocognitive models of schema effects in memory. It is possible that the congruency effects observed in Experiments 1 - 3 could be a consequence of retrieval-based strategies, in which a participant uses their prior knowledge of schema structure to deduce the correct category at the point of retrieval. An incongruency effect is, however, much harder to explain via such retrieval-based strategies. That said, there are a few retrieval-based strategies that could improve performance in the incongruent condition, if applied reliably. The following exploratory analyses aim to address two of these strategies and to elucidate whether they could be driving the incongruency advantage observed in Experiment 4.

Avoiding within-schema foils

The majority of Incongruent sequences occur in blocks in which two of the three sequences are schema-incongruent (the third being an Unrelated sequence). In addition to this, we also know from the post-experimental questionnaires that participants did tend to notice the incongruency manipulation. Given their explicit knowledge that some sequences change, it may be possible for participants to simply remember that a change occurred. Upon noticing this, a participant could simply apply the schema at retrieval and avoid any foil categories that are usually in the same schema as the cue. This would work because the correct incongruent category is guaranteed to come for the opposing schema. For example, suppose the schemas were: (red) *face-building-dog-car*, (blue) *furniture-tool-plant-insect*, and the participant saw an incongruent sequence: *face-building-tool-plant.* The participant would be cued with *building* and would need to retrieve *tool.* However, if the participant remembered that a change occurred, but not the specific category, they could deduce the correct category if the foil being used on that trial came from the same schema as the cue, in this case, *face, dog* or *car.* For example, if the foil category was *car,* the participant could eliminate that foil as the correct answer as an incongruent target can't come from the same schema as the cue.

To investigate this hypothesis I ran a 2x3 mixed-effects model with category-level Cue+1 accuracy as the dependent variable. The first predictor, Congruency, had three levels: Unrelated (baseline), Congruent, and Incongruent. The second predictor, Same-different, had two levels: Same and Different, referring to whether the target and the foil came from the same or different schemas. I ran this model using treatment contrasts and then broke the model down in a pairwise manner using *emmeans* with Holm correction for multiple comparisons.

Of particular note, there was no interaction between the Incongruent vs. Unrelated contrast and the Same-different predictor ($\beta = -0.13$, $SE = 0.18$, z = -0.74, p = .457) indicating that the effect of incongruency was equivalent in both the Same and Different condition. For completeness, I investigated the incongruency effect in each Same-different condition. When looking at trials in which the target and the foil came from different schemas, performance in the Incongruent condition was significantly higher than performance in the Unrelated baseline ($\beta = 0.44$, $SE = 0.07$, $z = 5.89$, p <.001). However, when investigating effects for trials where the target and foil came from the same schema, the difference between the Incongruent and Unrelated condition failed to reach statistical significance ($\beta = 0.31$, $SE = 0.16$, $z = 1.86$, $p = .063$). This suggests that there may be some evidence that the target and the foil coming from different schemas may be driving performance in the Incongruent condition.

However, it is worth inspecting the descriptive statistics here. In the Incongruent condition, when the target and the foil came from the same schema (i.e., when performance couldn't be enhanced by rejecting within-schema foils) performance was actually numerically higher

($M = 69.78\%$, $SD = 23.6\%$) than when the target and foil came from different schemas ($M = 68.15\%$, $SD = 12.03\%$). This finding is not in line with the interpretation that Incongruent performance was enhanced by the use of the retrieval-based strategy. Crucially, in the Unrelated baseline condition performance was higher when the target and the foil came from the same schema ($M = 63.33\%$, $SD = 14.26\%$) than when they came from different schemas ($M = 58.44\%$, $SD = 13.12\%$), and this difference was numerically larger than the difference in the Incongruent condition. This finding indicates that the lack of statistical significance in the Incongruent vs. Unrelated comparison when the target and foil came from the same schema is more so driven by increased performance in the Unrelated case, rather than performance being worse in the Incongruent condition.

Additionally, there is a caveat to consider here. When breaking the data down in this way there was a large imbalance in the number of trials in the Same and Different levels of the Same-different predictor. Only five (out of fifty) trials per participant had a target and a foil that came from the same schema at Cue+1 in the Incongruent condition. This imbalance resulted in a lack of statistical power which may explain why the Incongruent contrast narrowly missed significance ($p = .063$). This caveat highlights that this analysis should be interpreted with caution. Overall, the analysis does not seem to support the notion that the incongruency advantage was solely driven by the nature of the foils at test.

Avoiding foils that come first in the sequence

Another related strategy that a participant could use to enhance performance in the Incongruent condition would be to avoid foils that they know usually come at the beginning of a schema-congruent sequence. For example, if the Congruent schema was f*ace-building-dog-car* and the participant saw an Incongruent sequence *face-building-tool-plant*, the target category would be *tool*, but the foil category could be *face*. In such a case, the participant could eliminate *face* as the correct response because they know that this category was presented before the cue during encoding, not after.

To test whether this strategy could be driving performance in the Incongruent condition, I broke down the category-level data at Cue+1 based on whether the foil was the category that occurs at the beginning of the sequence. Crucially, I was interested in whether performance in the Incongruent condition differed when the foil was the first category in the Congruent schema. I found that accuracy did not significantly differ when the foil at Cue+1 was not the first category in the Congruent sequence (M = 68.03%, SD = 11.12%) as compared to when it was (M = 67.03%, SD = 22.08%), $t(45)$ = -0.37, p = .710, d = -0.06. This suggests that performance in the Incongruent condition was not driven by a retrieval-based strategy where participants eliminated foils based on the category being the first item in a congruent schema.

3.3.2.4.2 Prediction errors at encoding and retrieval-based performance

Experiment 4 was the only behavioural experiment in which I observed a significant incongruency advantage. However, all four experiments revealed significant evidence of prediction errors at the point of encoding. Accuracy and response times were consistently perturbed in the Incongruent condition at Cue+1, indicating an error regarding the category / animacy of the upcoming item. Interestingly, this suggests that the mere presence of a prediction error of some sort is not always sufficient to drive learning; rather, the specific nature of the prediction error may be of importance. I will return to this point in the general discussion. Importantly, this raises the question as to whether the behavioural errors that occurred at encoding bear any influence on accuracy for Incongruent trials at test. Here, I investigate whether performance at Cue+1 at encoding is predictive of performance in the Incongruent condition at test, and whether this depends on the kind of violation taking place (i.e., temporal reordering or categorical violation). To do this, I ran two mixed-effects models on the combined data from Experiments 3 and 4 with category-level retrieval accuracy at Cue+1 in the Incongruent condition as the dependent variable. The two models investigated whether encoding-based accuracy or encoding-based response time predicted retrieval-based performance respectively. The first between-subjects predictor, Violation type, had two levels: Sequential (Experiment 3) and Categorical (Experiment 4), and was included in both models. The second predictor varied across the two models. The accuracy model used encoding-based accuracy at Cue+1 as the predictor and had two levels: Accurate vs. Inaccurate, where inaccurate trials would indicate a prediction error. The second model used response time on accurate encoding-based Cue+1 trials as a continuous predictor, where longer response times are an indication of greater prediction error. These models were run using treatment contrasts and were then broken down using *emmeans* (accuracy model) or *emtrends* (response time model).

Accuracy model

The first model attempted to predict category-level accuracy in the Incongruent condition at Cue+1 during retrieval using accuracy at Cue+1 at encoding as a predictor. When collapsing across encoding accuracy (accurate vs. inaccurate), there was a significant effect of Violation type with performance being higher for categorical violations than for sequential violations ($\beta = -0.57$, $SE = 0.12$, z = -4.58, p <.001), as would be expected given that an incongruency advantage was only observed in Experiment 4. However, when breaking the model down by encoding accuracy, encoding-based performance did not predict performance at retrieval in either Experiment 3 ($\beta = -0.09$, $SE = 0.11$, z = -0.86, p = .392) or in Experiment 4 ($\beta = -0.1$, $SE = 0.11$, z = -0.9, p = .367). This indicates that the explicit behavioural errors at encoding were not the primary driver on any incongruency advantage at retrieval, but that this advantage was instead driven by the categorical vs. sequential distinction.

Response Time model

The second model predicted category level accuracy in the Incongruent condition at Cue+1 during retrieval using response time at Cue+1 at encoding as a predictor. Only trials where participants made an accurate animacy judgement at encoding were included in this analysis. When assessing the effect of Violation type at the mean encoding response time, I found that Violation type was a significant predictor of retrieval-based accuracy with performance being higher for categorical than sequential violations ($\beta = -0.56$, $SE = 0.13$, $z =$ -4.47, $p <.001$). Again, this is in line with our finding that only Experiment 4 produced an incongruency advantage. When breaking the model down by Experiment, I found that response time did not predict retrieval-based accuracy in Experiment 3 ($\beta = 0.16$, $SE = 0.41$, $z = 0.39$, $p = .698$) or in Experiment 4 ($\beta = 0.35$, $SE = 0.48$, $z = 0.73$, $p = .467$). This indicates that slower responses on Incongruent trials at encoding were not predictive of retrieval-based performance on those trials.

Overall, these analyses suggest the encoding-based performance was not predictive of category-level performance at retrieval on Incongruent trials. Importantly, this indicates that the incongruency advantage observed in Experiment 4 was probably not driven by violations at the animacy or motoric level and is instead more likely to be related to the type of violation, sequential vs. categorical.

3.3.2.5 Experiment 4 Discussion

Experiment 4 replicated the main findings from the previous experiments. At retrieval, there was a congruency advantage at the category, but not exemplar level. These effects were also preceded by evidence of ongoing predictions and prediction errors during sequence encoding. Overall, this suggests that across four behavioural experiments the congruency advantage at the category level, and the effects of Congruency on the animacy judgements at encoding, are very robust and, with minor exceptions, were consistently replicated across all experiments. Crucially, and in contrast to the previous experiments, Experiment 4 also revealed a significant incongruency effect at the category level with accuracy being greater in the Incongruent condition at Cue+1 relative to the Unrelated baseline. Notably, this incongruency advantage was not present at Cue+2 or at the exemplar level. This indicates that while violations of sequential order do not induce an incongruency advantage, violations of content do. I will review this finding in more detail in the Chapter's general discussion.

Exploratory analyses of Experiment 4 also revealed that the incongruency effect observed was unlikely to be a consequence of retrieval-based strategies. Two potential strategies were identified: (1) avoiding within-schema foils on Incongruent trials, because incongruency in Experiment 4 necessitates an out-of-schema category, and (2) avoiding foils that are usually the first item in the sequence, because this item is known to come first it can't be the correct answer. Exploratory analyses revealed that when these factors were controlled there was no impact on performance in the Incongruent condition, suggesting that these strategies were not responsible for the observed incongruency advantage and that encoding-based mechanisms were likely employed.

A final exploratory analysis investigated whether explicit behavioural errors during encoding predicted incongruent performance at test. Neither accuracy nor response time on the animacy judgements predicted Incongruent accuracy at retrieval. This indicates that the observed incongruency effect was not a consequence of explicit behavioural errors but was instead driven by the violation of the expected category.

3.3.3 Discussion

Across two experiments in this chapter, I observed that the type of violation was critical to the emergence of an incongruency advantage, with advantages only being observed for categorical and not sequential violations. This raises the question as to why violations in content, rather than violations in temporal order, would produce an incongruency effect? One obvious answer to this question is that the schemas generated in this paradigm simply don't contain information concerning temporal order, they only represent content. However, participants in Experiments 3 and 4 did seem to have explicit knowledge of sequence order. They were required to pass a test during each pre-learning phase, in which they reordered a scrambled sequence and wrote out the sequence using free recall. Moreover, when asked whether they noticed anything about the sequences in the post-experimental questionnaire, most participants in Experiment 3 ($n = 33$) and Experiment 4 ($n = 34$) stated that they noticed that the order or content was changed for some sequences. As such, it seems unlikely that the lack of an incongruency effect is due to a lack of knowledge regarding sequence order or a lack of awareness of violations to sequence order.

Another possibility is that it may be due to the nature of the predictions that are taking place. Rather than predicting which category will be occurring next, participants may have an ongoing predictive representation that contains information concerning all the categories that are expected in the sequence, regardless of whether a given category is expected to occur next or not. If this is the case, the change in temporal order may not trigger a strong prediction error because the incongruent category is still included in the representation of categories that are expected in that context. In contrast to this, a change in content, as in Experiment 4, could produce a stronger prediction error because the incongruent category is not included in the representation of currently expected categories.

However, one issue with this interpretation is that if temporal order information is absent from the schemas, and all categories are equally predicted, this would make incongruent sequences with sequential violations equivalent to schema-congruent sequences. If all categories are equally predicted at all times, order of presentation should not matter, and a schema congruency effect should be observed in the Incongruent condition. One solution to this problem is that the congruency advantage observed in the Congruent condition is purely driven by retrieval-based strategies. Such strategies would be ineffective in the Incongruent condition, even if all categories were equally predicted, explaining why no advantage was observed. However, exploratory analyses in Experiment 2 revealed that retrieval-based strategies may not fully account for the observed congruency advantages, suggesting that other processes may be at play.

The data here may be most in line with a model whereby predictions are occurring multiple steps into the future, with temporal order represented as the strength of activation of an upcoming category. With proximal categories represented more strongly than distal ones, temporal order could be represented, however, the partial activation of distal categories could be sufficient to inhibit strong prediction errors for sequential violations. Notably, congruency advantages would still be possible in a Congruent condition as the expected temporal order is still present.

In the following chapter I will use multivariate decoding to try and decode stimulus-specific representations before a category actually occurs. If the above explanation is correct, we would expect within-schema categories to be decodable (relative to out-of-schemas categories) throughout a sequence even if the category is not next.

3.3.3.1 Retrieval- vs. encoding-based accounts of schema effects at retrieval.

In Experiments 1 - 3, it was unclear whether encoding–based processes were contributing to the observed effects. I only observed congruency benefits, and it is possible that these effects could be driven by retrieval-based strategies. Some exploratory analyses suggested that retrieval–based strategies may not entirely account for these effects, but these analyses were by no means conclusive. The incongruency benefit observed in Experiment 4 is, however, much harder (though not impossible) to explain in terms of retrieval-based strategies. I have noted two retrieval-based strategies that might explain the incongruency effect. First, on incongruent blocks, it may be possible for participants to simply remember that a change occurred, while retaining no other instance-specific information. Knowing that the sequences changed, a participant could then apply the schema at retrieval but avoid any categories that are included in the same schema as the cue. Second, it may be possible to eliminate the foil category on Incongruent retrieval trials if that foil is normally the first item in the Congruent sequence. If the foil is known to be the first item in a Congruent sequence, it can't be the correct response. Although these techniques could in theory drive performance in the Incongruent condition, it is somewhat implausible that participants could reliably apply them while under the cognitive load of the task (assuming they recognised them as viable strategies to begin with). Moreover, the exploratory analyses provided little evidence that the incongruency effect observed in Experiment 4 was driven by these strategies. Overall, this suggests that encoding-based processes are likely contributing to the schema effects observed in Experiment 4, which is in line with neurocognitive models of how schema effects emerge.

I contend that the encoding-based mechanisms driving this incongruency effect relate to some kind of predictive coding process by which unexpected categories are triggering prediction errors that potentiate learning. However, it is worth considering other mechanisms that could be at work here. For example, we know from the post-experimental questionnaire that most participants notice the incongruency manipulation. If this is the case, participants might recognise that the unexpected items are of particular importance for the test phase. If so, rather than encoding being potentiated by prediction errors, participants may simply be engaging in explicit rehearsal of the incongruent items. Although possible, this rehearsal account may struggle to account for the lack of an incongruency effect at Cue+2. If the incongruency effect is due to prediction errors, the lack of an effect at Cue+2 is simple to explain: the prediction error at Cue+1 perturbs ongoing predictions, and therefore the schema-incongruency at Cue+2 does not produce a prediction error because it wasn't predicted to be schema-congruent. The rehearsal account struggles to account for this discrepancy, as it is unclear why Cue+1 would be rehearsable but Cue+2 would not. It might be argued that participants may be biased towards rehearsing the first incongruent item due to its proximity to the cue and due to issues surrounding cognitive load when rehearsing two items. However, this account is the less parsimonious explanation of this finding, which the predictive account can account for with fewer assumptions.

Finally, it may be argued that the lack of an incongruency effect in Experiments 1 – 3 was due to participants noticing the incongruency manipulation. If participants notice the manipulation, this may result in an explicit expectation of incongruent events, reducing surprise and any prediction errors that could drive learning. The finding of an incongruency effect in Experiment 4 makes this explanation less plausible. In both Experiments 3 and 4, post-experimental questionnaires revealed that the majority of participants noticed that some trials violated the pre-learned structure. However, despite explicit knowledge of the incongruency manipulation for both sequential and categorical violations, only categorical violations produced an incongruency advantage, indicating that explicit knowledge of the incongruency manipulation was not responsible for the absence of an incongruency advantage in Experiments 1 – 3.

3.3.3.2 Lack of an exemplar-level incongruency effect

One notable finding from Experiment 4, that is not in line with existing models of schema effects (e.g., SLIMM, van Kesteren, 2012), is the lack of an incongruency advantage at the exemplar level. SLIMM predicts that under schema-incongruent encoding conditions there should be enhanced memory for incidental details of the encoding instance. In the present experiment, this was expected to manifest as enhanced memory at the exemplar level in the incongruent condition, however this was not observed. It is worth noting that SLIMM predicts enhanced memory for *incidental* detail, i.e., detail that may not be task-relevant. In Experiment 4, this was not strictly the case, in fact, recalling exemplar-level detail was central to the task demands. It might be argued that the task-relevance of the exemplar-level detail in this task was the reason no schema effects were observed at the exemplar level. Indeed, Greve et al. (2019) observed an advantage for incidental detail in an incongruent condition, but this detail was completely irrelevant to the task at hand. However, it is unclear what the adaptive advantage of encoding incidental details, rather than task-relevant details, would be. If schemas have primarily evolved due to the advantage they provide in facilitating adaptive behaviours, one might expect that both incidental and task-relevant details would be equally important to encode when a surprising event occurred. When an unexpected event occurs, it is often unclear what factors caused this change or what factors might allow for its prediction in the future. As such, it would be most adaptive to encode all instance-specific information, incidental or not, such that a predictor could be deduced from its repeated occurrence in future incongruent events. Either way, the studies presented in this thesis have failed to provide evidence for enhanced memory for exemplar-level detail.

3.3.3.3 Prediction errors at encoding

Experiment 4 established that it is possible to observe incongruency effects when violations occur at the category but not sequential level. This distinction is notable because the encoding analyses indicated that prediction errors were present in all four experiments. That is, prediction errors were evident at encoding, but they did not consistently precede enhanced memory performance at test. This pattern suggests that the specific nature of a prediction error may be a critical determinant of its impact on learning. The present findings indicate that prediction errors during encoding likely occurred at the animacy level, yet this level of discrepancy was insufficient to facilitate memory improvements at the broader category level. Exploratory analyses showed that the presence or absence of an animacy violation had no effect on the emergence of an incongruency advantage at test, despite participants showing explicit behavioural errors to these animacy violations. Thus, the presence of a prediction error, in itself, does not appear to guarantee learning across all forms of task-relevant information.

Notably, prediction errors in this paradigm are indexed via a learned motor response. It is possible that this motor response is simply divorced from the underlying mechanism that is generating the incongruency effect in Experiment 4. Indeed, our exploratory analysis is in line with this notion, indicating that it was the categorical violation that was crucial for the emergence of the incongruency advantage. Nevertheless, the presence of reduced accuracy and increased response times at Cue+1 in the Incongruent condition is indicative of a prediction error taking place, but this error did not potentiate learning. As such, a key takeaway here may be that the prediction error may need to concern the specific information that is to be recalled, hence why categorical violations produced a category-level incongruency advantage (they are the same level of information), but violation in animacy and/or sequential order did not. Likewise, this may also shed light on our lack of an effect at the exemplar level, simply, violations were not occurring at the exemplar level.

3.3.3.4 Conclusion

In this chapter I have successfully replicated the U-shaped function predicted by SLIMM, however, under limited circumstances. In the present paradigm, it appears this effect is only present when violations occur at the category level, with sequential violations being insufficient to produce a memory advantage at the Incongruent side of the curve.

In the following chapter I combine Experiment 4's behavioural paradigm with magnetoencephalography (MEG). This Experiment has three broad goals. First, I aimed to replicate the behavioural findings of Experiment 4 to ensure the reliability of the incongruency effect. Second, I attempted to provide evidence of stimulus-specific neural predictions prior to a category's actual occurrence. I did this by applying multivariate decoding techniques to the pre-stimulus period. Using variables such as decoding accuracy and the oscillatory power spectrum of decoding time series, I aimed to uncover evidence of ongoing predictive processing during sequence encoding. By breaking this data down by whether a category is expected to occur next, or whether it is a within or out-of-schema category, I could also be able to shed light on the nature of the predictions taking place (do they concern a single category, or multiple upcoming categories). Third, I investigated how neural responses differed between the Congruency conditions post-stimulus onset. To do this, I contrast time-frequency responses across the experimental conditions to assess changes in spectral power dependent on whether a schematic prediction was affirmed or violated. In doing so, I provide evidence that encoding-based mechanisms are indeed contributing to the effects observed at retrieval.

4 The Neural Dynamics of Schema Effects via Magnetoencephalography

4.1 Introduction

In this chapter, I will investigate the neural mechanisms underlying the effects observed in the previous behavioural experiments. To do this, I will combine the behavioural paradigm from Experiment 4, where I observed a U-shaped function of schema congruency at the category level, with magnetoencephalography (MEG) and multivariate pattern analysis (MVPA). Specifically, I will examine whether the emergence of schema effects is preceded by evidence of ongoing neural predictions during the pre-stimulus period, and whether the affirmation or violation of these predictions is followed by prediction error signals.

Before presenting the experiments in this chapter, I will briefly review three relevant topics: (1) the neuroscience of schema processing, (2) the use of MVPA to study neural representations, and (3) previous research on pre-stimulus neural predictions.

4.1.1 The neuroscience of schemas

Numerous brain regions have been implicated in schema processing. Most prominent among these are the hippocampus (HPC) and the medial prefrontal cortex (mPFC). However, other regions such as the anterior temporal lobe, the middle and superior temporal gyri, the temporoparietal junction, and the angular gyrus have also been associated with schema-related functions (Gilboa & Marlatte, 2017). This distributed collection of brain regions aligns well with the mnemonic, semantic, and attentional networks that one might expect to be involved in the processing of schemas (Buckner et al., 2008; Corbetta & Shulman, 2002; Jackson, 2021; Kim, 2010; Ralph et al., 2016; Thakral et al., 2017). Given that schemas are believed to act as a template against which ongoing experiences are routinely interpreted, it is of little surprise that neuroimaging research would reveal engagement of regions associated with episodic recollection, semantic control, and stimulus-driven attention and saliency signals.

Although there is a general consensus that the hippocampus and medial prefrontal cortex are crucial for schema processing, the precise contributions of each region are still debated. SLIMM (van Kesteren et al., 2012) predicts a division of labour between the mPFC and HPC with either region being specifically involved in schema-congruent or schema-incongruent encoding respectively. Under this framework, the mPFC enacts top-down inhibition of HPC encoding when events are schema-congruent, resulting in gist-like encoding, whereas the HPC is particularly sensitive to prediction errors and supports enhanced encoding of episodic detail. Indeed, several studies have supported this proposal of an mPFC-HPC distinction. van Kesteren et al. (2013) had participants study scene-object pairs during fMRI and rate them on how likely they were to co-occur in the real world. The following day, recognition and associative memory for the scene-object pairs was highest for pairs rated as most likely to co-occur. Crucially, mPFC activity increased linearly with subjective ratings of congruency, whereas parahippocampal areas of the medial temporal lobes showed increasing activity with decreasing congruency. Together, these findings are in line with SLIMM's prediction that congruent information is detected by the mPFC which subsequently inhibits activity in the medial temporal lobes. In a different study, it was found that patients with mPFC lesions show fewer false alarms to new items that are perceptually similar to previously viewed items, potentially indicating greater pattern separation due to less inhibition of the HPC by the mPFC under schema-congruent conditions (Spalding et al., 2015). SLIMM has also been corroborated by its prediction of a U-shaped function of schema congruency that has now been demonstrated across several studies (Greve et al., 2019; Ortiz-Tudela et al., 2024; Quent et al., 2022; Yacoby et al., 2025), including by Experiment 4 in the preceding chapter. However, as has been revealed both in this thesis and in the wider literature (Ortiz-Tudela et al., 2024; Yacoby et al., 2025), the presence of this U-shaped function appears to be highly context-dependent.

Beyond the fact that the U-shaped function is not observed as consistently as SLIMM might predict, several neuroimaging studies suggest that mPFC-HPC interactions may be more complementary than proposed by SLIMM. Raykov et al. (2019) familiarised participants with the lead character's of a TV show over the course of a week to create schematic knowledge regarding the characters identities. At test, a schema congruency effect was observed where participants had better recognition memory for novel pictures of the characters they had been trained on over those they had not. Critically, whole-brain fMRI analyses revealed that both the ventral mPFC (vmPFC) and the HPC showed greater activation for trained vs. untrained identities, with no interaction between the regions, suggesting that both regions had a similar preference for schema-congruent information. Moreover, trained vs. untrained identities did not modulate vmPFC-HPC connectivity. These findings are at odds with SLIMM that would predict greater preference for schema congruency in the mPFC, and for schema congruency to modulate mPFC-HPC interactions. In a different study investigating expectation violations, Garrido et al. (2015) investigated the neural dynamics of prediction violations by combining dynamic causal modelling (DCM) with MEG. They found theta activity in both the vmPFC and HPC was engaged by prediction violations. Crucially, DCM on theta phase-locked activity suggested that violation-induced theta in the HPC was driven by vmPFC theta, a finding not predicted by SLIMM. Taken together, these studies indicate that both schema-congruent and schema–incongruent events engage the HPC and mPFC, suggesting that neither region is predominantly selective for only one type of information.

Despite an ongoing debate regarding the precise contributions of the HPC and mPFC to schema effects in memory, there is broad consensus that schema effects are the consequence of a predictive coding process (Gilboa & Marlatte, 2017; Huang et al., 2023; Ortiz-Tudela et al., 2024; Quent et al., 2021; van Kesteren et al., 2012; van Kesteren & Meeter, 2020). Under this framework, schemas facilitate ongoing neural predictions that are continuously compared with sensory inputs. This predictive processing modulates mPFC-HPC interaction which results in the behavioural manifestation of schema effects. In line with this predictive account, van Kesteren et al. (2012) note that schema effects can be modeled using a Bayesian framework, where a prior distribution represents existing schematic knowledge and a likelihood distribution represents incoming sensory experience. Indeed, the lack of an incongruency effect in an Unrelated condition, despite maximal novelty, is well-explained from a Bayesian perspective. When both the prior and likelihood distributions are both relatively flat (i.e., inputs are novel but predictions are imprecise), the lack of a strong prediction leads to little prediction error, and little learning occurs. More concretely, behavioural evidence of ongoing predictions has also been linked to enhanced memory. In a study where participants were trained on a board game across multiple sessions, schema-based predictive eye movements during encoding were associated with enhanced memory retention, effects that were most prevalent in participants with the most well-developed expertise (Huang et al., 2023). Finally, neuroimaging studies of prediction and learning frequently report neural correlates of prediction error in regions associated with schema-related processes (e.g., mPFC-HPC). In particular, hippocampal theta oscillations and PFC-MTL connectivity are frequently linked to prediction violations (Garrido et al., 2015; Recasens et al., 2018), indicating that these regions support the comparison of incoming events with prior expectations.

Despite the evidence that schema effects are driven by predictive processing, the precise temporal dynamics by which this occurs remains underexplored. In particular, while prediction violation and the pre-stimulus activation of predictable representations have been observed independently, they have not yet been jointly linked to schema effects in a single study. The present study seeks to bridge this gap by using multivariate pattern analysis (MVPA) to test whether schema-based expectations manifest as pre-stimulus neural reactivation, and whether these anticipatory signals are subsequently followed by prediction error responses and behavioural schema effects. Before doing this, I will first move on to introduce the use of MVPA in neuroscience more generally.

4.1.2 Multivariate pattern analysis in neuroscience

The use of multivariate methods in cognitive neuroscience first came to prominence via the seminal fMRI paper of Haxby et al. (2001). This research allowed for the discrimination of neural representations that classic univariate approaches could not. Haxby et al. (2001) demonstrated that patterns of activity across voxels in the ventral temporal cortex could distinguish between object and face representations, even when average activation did not significantly differ. This finding highlighted that stimuli are often represented in terms of distributed patterns of activation rather than in the activation of distinct regions or voxels.

Crucially, the introduction of these methods has allowed for the decoding of task-relevant information both when a stimulus is online but also during periods of spontaneous neural activity when that information may still be cognitively relevant (Liu et al., 2022; Staresina et al., 2013). A classic example of decoding spontaneous activity comes from episodic memory research where the representation of a memory can be decoded from its replay during recall or consolidation. For example, Deuker et al. (2013) combined Support Vector Machines (SVM) with concurrent fMRI and EEG. Participants learned object-location paired-associates and SVM classifiers were trained on the accompanying neural data. In a subsequent rest period, participants slept in the scanner. It was found that stimulus-specific representations could be decoded during the resting period both during sleep and during wakeful rest. Thus, multivariate methods have provided a method for analysing brain data and decoding stimulus-specific representations when the averaging of event-related responses is not valid or possible.

The extension of multivariate decoding techniques to high-precision time series data such as MEG has further highlighted the approach's utility in advancing our understanding of human cognition. Owing to MEG's superior temporal resolution compared to fMRI, it is possible to decode representations with millisecond precision (Grootswagers et al., 2017). This makes it possible to pinpoint precisely when a representation emerges and how it evolves over time. A prominent example of this comes from the method known as time generalisation (King & Dehaene, 2014). By training and testing classifiers on every combination of training and test time points, it can be established how well a pattern of activity at one time point predicts the pattern of activity at another, giving an indication as to whether a representation is stable across time or whether it transforms as cognition unfolds.

Importantly, decoding methods have also allowed researchers to establish the temporal dynamics of visual perception and precisely when different forms of visual information are decodable from a brain signal (T. Carlson et al., 2013). Features such as category, animacy and exemplar information all have different time courses with respect to the onset and the peak of their decodability. T. Carlson et al. (2013) found that category membership could be decoded from a neural signal as early as 80 ms post stimulus onset with peak decoding occurring anywhere between 120 - 180 ms. However, higher-level conceptual characteristics such as animate vs. inanimate distinctions did not become apparent until 160 ms with a peak decoding time of 240 ms. Notably, the discriminability of a set of stimuli can be dependent on the way in which they are grouped. For example, the same set of stimuli could be discriminated earlier when grouped as Faces vs. Bodies than when grouped as Human vs. Animal, further emphasising that more abstract information emerges later in the processing hierarchy.

Importantly, MPVA has also been used to reveal the dynamics of memory retrieval. By applying MVPA both at encoding and retrieval, it has been possible to determine specific temporal windows that are associated with the reactivation of memory traces (Jafarpour et al., 2013). Early decoding during stimulus encoding has been associated with enhanced memory at test. Moreover, during cued recall, it is possible to decode associates of the cue from approximately 500 ms after cue onset, with successful decoding predicting accurate recall (Bramão & Johansson, 2018). Crucially, decoding also seems to be possible during inter-stimulus or pre-stimulus periods. For example, in a working memory task, Fuentemilla et al. (2010) demonstrated stimulus decoding during a maintenance period whereby neural patterns representing specific items were periodically reactivated in synchrony with the phase of ongoing theta cycles. Likewise, Moccia et al. (2022) found that a retrieval goal (retrieval of visual vs. auditory targets) could be decoded from a pre-cue EEG time series, indicating that the ongoing retrieval context was represented in the pre-stimulus period. Together, these studies indicate that task-relevant representations can be reinstated or maintained in memory during periods when a stimulus is not actively perceived, and that MVPA provides a powerful tool for revealing their presence.

In the present chapter, I will use MVPA in an attempt to detect the predictive reactivation of upcoming categories in a sequence. By using MEG's fine temporal resolution, we can pinpoint the peak decoding of a category and attempt to generalise its sensor space representation to pre-stimulus periods when the category is expected but not currently online. In a preliminary pilot experiment, I first examine whether the behavioural task performed during classifier training influences classifier fidelity. Previous research suggests that engaging in a task, rather than passively viewing images, bears little influence on classifier performance (T. A. Carlson et al., 2003). Nonetheless, decoding representations during pre-stimulus periods likely entails poor signal-to-noise ratios, so investigating potential avenues of enhancing classifier fidelity is still warranted.

4.1.3 Pre-stimulus markers of prediction and memory

Before moving onto the Experimental work in this chapter, it is worth reviewing previous research that has revealed pre-stimulus markers of prediction and memory in the brain. For example, Gilboa & Moscovitch (2017) used electroencephalography (EEG) to provide evidence of pre-stimulus schema monitoring in the mPFC. Participants viewed faces of familiar people who they personally knew, famous people they did not personally know, and unfamiliar faces, and were required to indicate whether they had personally met the person before. Here, the self acted as a schema against which incoming information could be assessed. In the 500 ms period prior to a face, theta coherence desynchronisation was observed between the mPFC and regions of the inferior and lateral temporal cortices, areas known to be involved in face processing (Kanwisher et al., 1997; Sergent et al., 1992). This pre-stimulus desynchronisation predicted behavioural accuracy, and modulated the N170 ERP and a frontal positive ERP at approximately 200 ms, components known to be associated with face processing (Itier, 2004) and semantic organisation (Blanchet et al., 2007) respectively. Gilboa & Moscovitch (2017) interpreted these results as reflecting ongoing schema instantiation by the mPFC, whereby the region maintains a context-relevant schema to support the processing of upcoming stimuli, facilitating memory retrieval and cognitive performance. However, while this study demonstrated that the mPFC can maintain schema-relevant information prior to stimulus onset, it did not provide evidence of schema-based modulation of sequential predictions or the pre-activation of specific upcoming events. Thus, the extent to which pre-stimulus neural activity reflects not only general schema instantiation but also the predictive reactivation of upcoming items in a sequence remains unclear.

More recently, research has gone beyond general schema instantiation to reveal evidence of category-specific pre-stimulus activity, detectable via multivariate analyses. Sherman et al. (2022) combined intracranial EEG (iEEG) with a statistical learning task in which a continuous sequence of scene categories had a predictable structure. Using a representational similarity analysis (RSA) approach on sensor-level time-frequency data, it was shown that patterns of activity prior to a category's onset were significantly correlated with the categories online frequency-by-sensor representation. Sherman et al. (2022) related these effects to subsequent memory and found that neural similarity evidence for a category prior to its occurrence was greater for remembered items than it was for forgotten items. This indicates that not only can representations of predictable events be brought online prior to their occurrence, but that this pre-activation benefits subsequent memory, a finding which is in line with the predictive account of schema effects in memory. Notably, however, Sherman et al. (2022) did not manipulate prediction violations (i.e., all predictions were affirmed). Moreover, predictions in this study arose in a single task in which a continuous stream of exemplars followed a statistical structure, allowing participants to anticipate repeating patterns. In contrast to this, schemas are thought to be long-term semantic representations that are formed across multiple episodes and then generalised to novel situations. It therefore remains unclear whether the predictive effects observed here would emerge in paradigms such as the one used in this thesis, where schema knowledge is acquired a priori and then applied to new events.

Finally, research shows that predictive cues can elicit the pre-stimulus activation of sensory templates. Kok et al. (2017) combined MEG with multivariate decoding and a paradigm where participants heard predictive audio cues that predicted the orientation of an upcoming grating. Multivariate classifiers applied to the pre-stimulus period between the cue and the target revealed that orientation could be decoded up to 40 ms prior to stimulus onset. Notably, a subsequent study revealed that predictive signals may also manifest in the spectral characteristics of the decoding time series. Hetenyi et al. (2024) conducted a similar study that attempted to decode shape categories rather than gratings. Classifiers were trained for these categories while the category was visually perceived and then applied to the time series between the auditory cue and the onset of the visual category. A Fast

Fourier Transform (FFT) was then applied to the decoding time series to investigate its spectral characteristics. It was found that the pre-stimulus classification time series for predictable categories contained more alpha power (10 - 11 Hz) than decoding time series from classifiers trained on randomly labeled data. Moreover, pre-stimulus alpha power correlated with post-stimulus shape discrimination. When an auditory cue was a valid predictor of shape category, pre-stimulus alpha power in the decoding time series predicted shape discrimination more so than on trials where the cue was invalid. Again, these results indicate that predictable information can be pre-activated prior to an event's actual occurrence, and that this pre-activation can influence perceptual performance. Notably, however, this study did not assess subsequent memory. As such, it remains unclear as to how these predictions, and their affirmation or violation, relate to memory retention.

In the coming sections I aim to investigate whether the schema effects observed in the previous behavioural experiments are accompanied by evidence of predictive processing in the brain. To do this, I combined the behavioural paradigm from Experiment 4 with multivariate decoding and MEG. In a preliminary pilot Experiment, I investigate whether classifier accuracy is influenced by the task performed during classifier training: Oddball, Category discrimination, or Exemplar discrimination. The results from the pilot were used to inform the design of a localiser task that would be used to train the category-specific classifiers in the main MEG Experiment.

In the main MEG experiment, I opted to use the behavioural paradigm from Experiment 4, as this was the only experiment in which the predicted U-shaped function of schema congruency was observed (van Kesteren et al., 2012). Participants performed the main temporal order memory task during MEG, followed by the localiser task determined from the pilot. Using this data, I conducted time-frequency decomposition to investigate whether Incongruent trials elicited neural violation signals, such as increases in theta-band activity, that have previously been associated with prediction errors in the brain (Garrido et al., 2015; Recasens et al., 2018). Following this, I trained category-specific classifiers on the localiser data and applied them to the time series during the main temporal order memory task. By examining classifier performance and the frequency characteristics of the decoding time series during the pre-stimulus period, I aimed to determine whether category-specific representations were activated in advance of a category's actual occurrence, consistent with schema-facilitated predictions.

The present study had three main aims:

1. Replicate the behavioural schema effects observed in Experiment 4, where pre-learned schematic knowledge of sequence structure produced a U-shaped function of schema congruency at the category level.
2. Link category-level incongruency effects to prediction error signals, such as theta-band responses.
3. Investigate whether these schema effects and prediction error signals are preceded by pre-stimulus predictive reactivation of expected categories.

4.2 Classifier Piloting

In this pilot Experiment, I aimed to investigate whether the task used during classifier training bears any influence on subsequent classifier fidelity. Previous research suggests that the task used during classifier training has little influence on subsequent performance (T. A. Carlson et al., 2003). However, given that the main MEG experiment presented in this chapter investigates pre-stimulus activity where SNR will be poor, even small enhancements in classifier fidelity would be beneficial. This pilot also offered the opportunity to ensure that the category images used in the previous experiments could indeed be distinguished using multivariate decoding.

During MEG, three different tasks required participants to evaluate the categories at different levels of representation. An oddball task was used as a baseline in which participants passively viewed the categories without any explicit instruction to retain category or exemplar-level information. This was accompanied with a category discrimination and an exemplar discrimination task, which required participants to encode information at either the category or exemplar level respectively. Classifiers were subsequently trained on the oscillatory activity associated with each category using a cross-validation approach. By comparing classifier performance across tasks, I aimed to reveal (1) whether the categories could indeed be decoded from an MEG time series, and (2) whether the overall performance of the classifier varied depending on the task used during training.

4.2.1 Methods

4.2.1.1 Participants

Nine participants were recruited from the University of York student body and were reimbursed with course credit. One participant was excluded because they did not perform above chance in the behavioural task as determined by a one-tailed binomial test performed on their trial-level data. This left a final sample size of $N = 8$ participants (5 females, 3 males; $M \pm SD$ age in years = 22.9 ± 2.82). All participants indicated that they were fluent English speakers, right-handed, had normal or corrected-to-normal vision, had no prior or existing neurological or psychological illnesses, and were not on any psycho-active drugs. Participants gave informed consent for the experiment, which was approved by the York Neuroimaging Centre's (YNiC) ethics committee.

4.2.1.2 Materials

This experiment used the same exemplar images from the eight categories used in the previous behavioural experiments: faces, dogs, insects, plants, buildings, cars, furniture, and tools. See Chapter 1 Methods for details.

4.2.1.3 Design

This experiment had a one-way design with a single within-subject independent variable, Task, that had three levels: Oddball, Category discrimination, and Exemplar discrimination. The dependent variable was accuracy coded as a binary variable.

4.2.1.4 Procedure

A schematic of the experimental procedure can be seen in Figure 4.1. Participants engaged in three tasks: Oddball, Category Discrimination, and Exemplar Discrimination. Each task was completed back-to-back during consecutive MEG runs. The order of the tasks was counterbalanced across participants.

Each task consisted of 14 blocks. In each block, participants would see a continuous stream of exemplars from the eight categories. Each block consisted of four exemplars from each of the eight categories for a total of 32 trials per block. Each stimulus was shown on the screen for 1000 ms with an inter-stimulus interval (ISI) jittered between 400 - 600 ms for an average ISI of 500 ms. All 14 blocks contained the same four exemplars from each category, however, the order of the exemplars in each block was randomised. Different exemplars were used in each task. This gave a total of 56 presentations of each of the eight categories per task giving 448 stimulus presentations per task.

4.2.1.4.1 Oddball task

The oddball task was intended as a baseline condition in which the participants passively viewed the images, and it was not necessary to process the stimuli at either the category or the exemplar level. In the oddball task, there were an additional four trials per block (giving a total of 36 trials per block) that acted as oddball trials on which the participants must respond. On these oddball trials, participants would see an exemplar in the same way as any other trial except that a red dot would be superimposed somewhere on the image. Participants were required to indicate when a trial contained a red dot by pressing a button on a response pad. Oddball trials were removed from the data prior to MEG analysis.

4.2.1.4.2 Category discrimination task

The category discrimination task was used to assess whether classifier performance improved when participants were required to process the stimuli at the category level, but not exemplar level. In the category discrimination task, four trials per block were followed by a two-alternative forced-choice test. After a trial, a visual mask would appear on the screen for 500 ms to indicate that a test was about to occur. Two categories would appear, one on the left and one on the right-hand side of the screen. One category would be a different exemplar of the same category that the participant just saw, the other would be an exemplar of a different category. The participant had to select the exemplar that was of the same category as the trial they just saw by pressing a button on a response pad. Participants were given 2000 ms to respond before they were timed out. If they were timed out the words “TIME OUT!” would appear on the screen in red text for 1500 ms. Time out trials were marked as incorrect. There were four tests per block that occurred after randomly selected trials.

4.2.1.4.3 Exemplar discrimination task

The exemplar discrimination task required the participants to process the stimuli at the exemplar level. The exemplar discrimination task was identical to the category discrimination task apart from the nature of the exemplars used during tests. During tests, participants would see two exemplars of the same category. One would be the exact same exemplar as the one they just saw on the previous trial, the other would be a different exemplar from the same category. Participants were required to select the exemplar that was the same as the one they just saw on the previous trial.

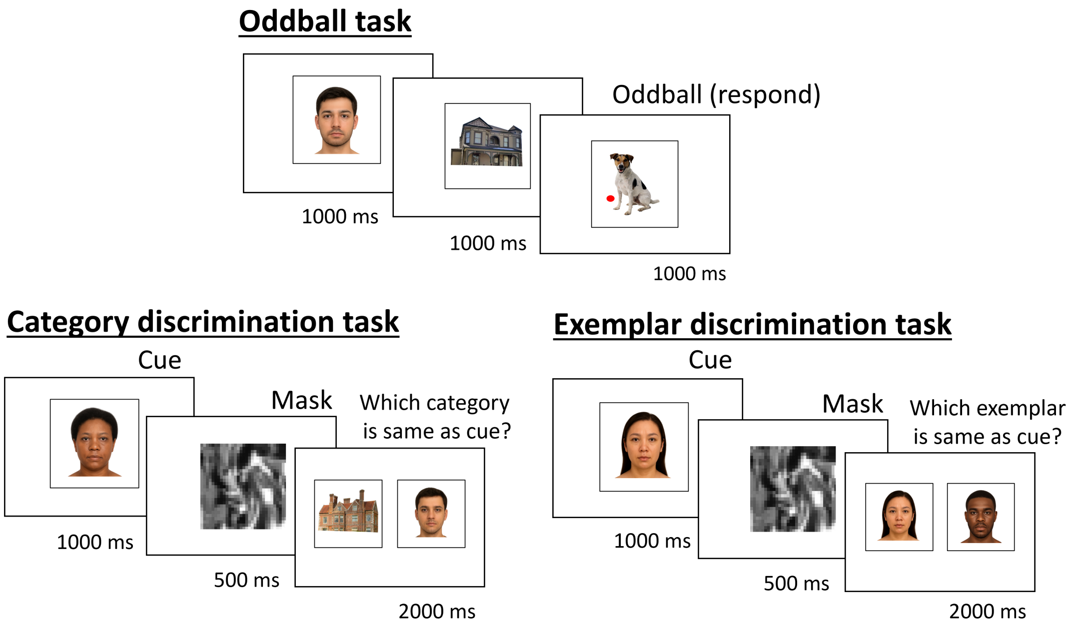

Figure 4.1: Schematic of the task used in the MEG pilot. During the oddball task participants saw a continuous stream of exemplars, 36 per block. Four of those trials were oddballs containing a red dot. Participants were instructed to respond whenever they saw an image containing a red dot. During the category discrimination task participants saw a continuous stream of exemplars, 32 per block. After four randomly selected trials per block, participants made a two-alternative forced-choice discrimination judgement as to which category was the same as the one just seen. During the exemplar discrimination task, participants saw a continuous stream of exemplars, 32 per block. After four randomly selected trials per block, participants made a two-alternative forced-choice discrimination judgement as to which exemplar was the same as the one just seen.

4.2.1.5 MEG data acquisition and processing

MEG recordings were acquired using a 4D Neuroimaging Magnes 3600 whole-head 248-sensor magnetometer system. Four of the sensors were offline giving a total of 244 sensors in the final data set. Data was collected at a sampling rate of 1000 Hz. All analyses were performed in sensor space, so no head digitisation was performed. Data preprocessing and analysis were performed using the MATLAB toolbox FieldTrip (Oostenveld et al., 2011).

Raw MEG recordings were first bandpass filtered between 0.5 - 45 Hz using Butterworth filters and notch filters were applied at 50 and 60 Hz to remove any residual noise from the mains frequency and projector frame rate. Bad channels were rejected via visual inspection. The data were then subjected to Independent Components Analysis (ICA) which was used for the removal of eye blink and cardiac artifacts. Finally, the data were epoched into -400 ms to +1000 ms segments around stimulus onset. Note that no trial rejection was performed in any of the MEG work presented in this chapter. The decision not to reject bad trials was motivated by the fact that in the main temporal order memory experiment, individual epochs form a part of larger sequences and rejecting individual trials would remove data from these ongoing sequences. Moreover, this decision was also justified by recent findings suggesting that the increased SNR that trial rejection provides does not compensate for loss in statistical power (Delorme, 2023).

4.2.1.6 MVPA

MVPA analyses were conducted using the MVPA-Light framework (Treder, 2020). All models used Linear Discriminant Analysis (LDA) to find the lower-dimensional projection of the data that maximised class separability. The analyses used a 5-fold cross-validation procedure with ten repetitions. A classifier was trained for each of the eight categories (faces, dogs, insects, plants, buildings, cars, furniture, and tools) using a one-vs-all approach, resulting in an 87.5% class imbalance in favour of the negative class. To account for this imbalance, classifier performance was quantified using Area Under the Receiver Operating Characteristic Curve (ROC AUC). Classifiers were trained on labeled MEG epochs, with sensors as features. Prior to training, each feature was z-scored across epochs per timepoint. MVPA-Light performed time-resolved decoding by applying LDA on a timepoint-by-timepoint basis (e.g., train on -400ms and test on -400ms etc.), yielding an

ROC AUC metric for each time point, category, and task.

4.2.2 Results

4.2.2.1 Behavioural

Given the small sample size and the exploratory nature of this pilot study, I did not perform inferential statistics on the group behavioural data. Behavioural accuracy in the three tasks is plotted in Figure 4.2. Average accuracy was above 90% in all tasks. However, there was some indication of worse performance in the category discrimination task relative to the oddball and exemplar discrimination tasks. Nonetheless, all participants were performing well above chance (50%) indicating that they were engaging with the task as expected.

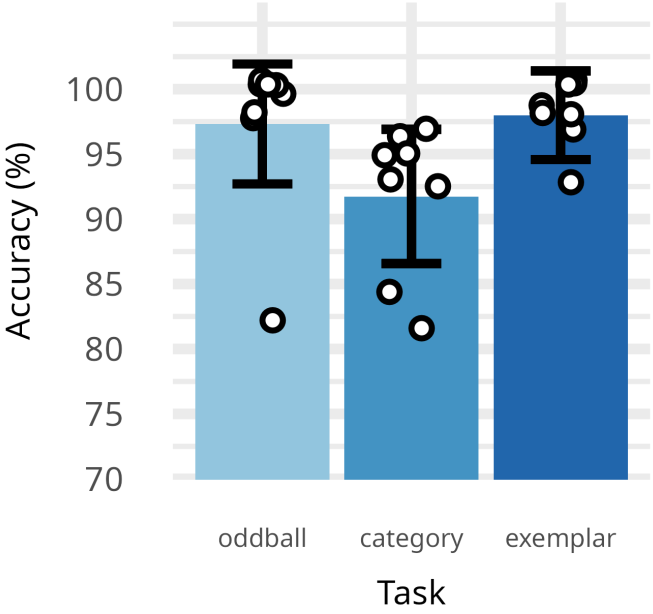

Figure 4.2: Mean accuracy (%) in each of the three tasks used in the MEG pilot experiment. Dots represent individual participant means. Error bars represent 95% confidence intervals.

4.2.2.2 MVPA results

The results from the pilot MVPA analysis are plotted in Figure 4.3. The first analysis investigated category decoding across the three tasks when the classifiers were trained on the full sampling rate of 1000 Hz. Category decoding peaked at approximately 200 ms post stimulus onset with little difference in the classifiers' peaks across the Oddball (M = 0.61, SD = 0.06), Category discrimination (M = 0.6, SD = 0.03), or Exemplar discrimination task (M = 0.6, SD = 0.03). Classifier performance began to rise at approximately 80 ms post-stimulus and returned to baseline by 600 ms in all three tasks.

I next ran the same analysis investigating whether averaging the MEG data across time could boost classifier performance. If a representation is maintained in the brain across time, averaging across time points may help to reduce noise that is unrelated to the ongoing category representation and boost classifier performance. The data at a sampling rate of 1000 Hz was averaged across 20 ms bins to give a final sample rate of 50 Hz. In this analysis, classifier performance again peaked at approximately 200 ms and there was a small numerical increase in classifier performance in the Oddball (M = 0.63, SD = 0.06), Category discrimination (M = 0.62, SD = 0.03), and Exemplar discrimination task (M = 0.61, SD = 0.03). The period of decoding was also similar to the previous analysis with classifier performance beginning to rise at approximately 80 ms post-stimulus and returning to baseline by approximately 600 ms in all tasks.

Overall, the Oddball task appeared to result in the highest numerical classifier fidelity. However, the Oddball task also had the widest error margins, that overlapped substantially with the performance of both the Category and Exemplar discrimination task classifiers, making the difference in performance difficult to dissociate from chance.

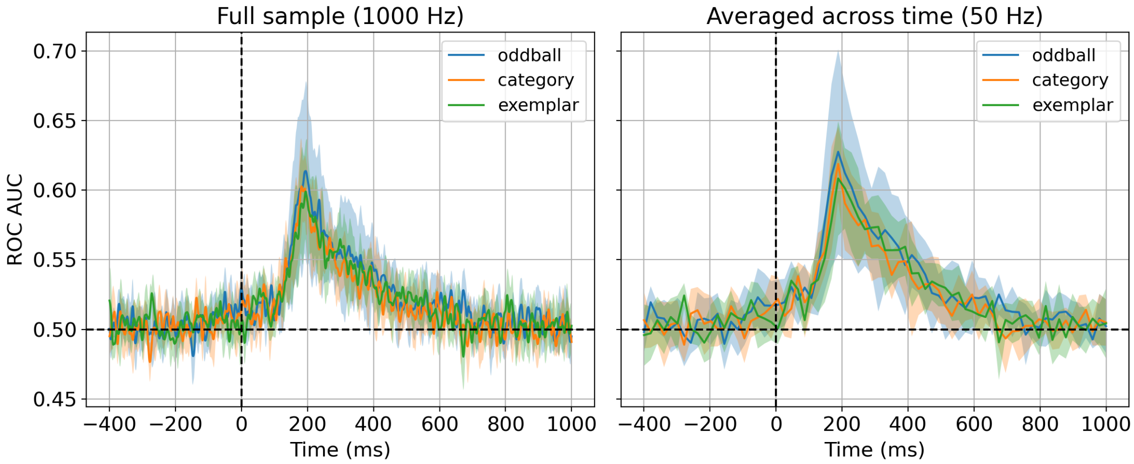

Figure 4.3: Classifier performance in the MEG pilot experiment from -400 to 1000 ms in each of the three tasks when performing on the full sampling rate of 1000 Hz (Left) and when the data were averaged across time in 20 ms bins to produce a rate of 50 Hz (Right). Shaded areas represent 95% confidence intervals.

4.2.3 Discussion

This pilot experiment demonstrates that the exemplars in the present dataset are decodable at the category level from sensor space MEG signals. However, consistent with previous research (T. A. Carlson et al., 2003), there was little evidence that classification accuracy was impacted by the task performed during classifier training. This indicates that the choice of a task used as a localiser is unlikely to be of substantial consequence. This lack of a distinction between tasks, and the early rise (~100 ms) and peak (~200 ms) of the decoding time series, also suggests that the categories are primarily being decoded using early visual information, and that category discrimination is less likely to be based on higher-level conceptual characteristics.

Having confirmed that the categories are indeed decodable, I moved on to the main MEG experiment. In the main MEG experiment, classifiers were trained using the exemplar discrimination task. The exemplar discrimination task was chosen as it showed greater behavioural accuracy than the category discrimination task and there was less error associated with its peak decoding estimate when compared with the oddball task.

4.3 MEG Experiment

The main MEG experiment investigated the neural correlates of the schema effects observed in Experiment 4. As before, participants were pre-trained on the sequence schemas before completing the main temporal order memory task during MEG. The key question was whether the schema effects observed in previous experiments were accompanied by evidence of predictive processing.

To test this, I applied time-frequency decomposition to Cue, Cue+1 and Cue+2 at encoding across the three Congruency conditions, examining whether prediction error signals and markers of memory encoding were reflected in modulations of theta-band activity. Following this, multivariate decoding was used to investigate decoding in the pre-stimulus period. If significant decoding could be revealed in the pre-stimulus time series, this would provide strong evidence that prior knowledge was used to generate predictions about upcoming events. Time-frequency decomposition could then show that these predictions are followed by prediction error signals, and the behavioural manifestation of mnemonic schema effects.

4.3.1 Methods

4.3.1.0.1 Participants

Due to the use of MEG in this experiment, a reduction in the number of trials was required to ensure participants were not in the scanner for more than an hour. To determine an appropriate number of trials and sample size, a Monte Carlo resampling procedure with

1000 iterations was conducted on the data from Experiment 4. For 30 participants with 27 trials per condition, the observed data yielded 90.1% detection of the schema incongruency effect and 100% detection of the schema congruency effect at the category level at Cue+1. As such, I aimed for a sample of at least $N = 30$ with 27 trials per condition. Based on the time constraints of the project a total of $N = 38$ participants were run to keep the sample as close in size as possible to the behavioural experiments.

Thirty-eight participants were recruited from the University of York student population. Participants were reimbursed with course credit or a £20 Amazon voucher. One participant was excluded due to a software error that meant their MEG data could not be recorded, however, their pre-study data was still included in the pre-learning analysis. This left a final sample size of $N = 37$ participants (31 females, 6 males; $M \pm SD$ age in years = 21 ± 3.02). All participants indicated that they were fluent English speakers, right-handed, had normal or corrected-to-normal vision, had no prior or existing neurological or psychological illnesses, and were not on any psycho-active drugs. Participants gave informed consent for the experiment, which was approved by the York Neuroimaging Centre's (YNiC) ethics committee.

4.3.1.0.2 Materials and procedure

The study was split into two sessions separated by approximately 24 hours. During the first session, participants engaged in the same pre-learning task used in the previous behavioural experiments. In contrast to the previous experiments, I extended the inter-stimulus interval between the items in a sequence from 500 ms to 1000 ms. This was done to increase the amount of time in which pre-stimulus activity could be decoded. After the 24-hour delay, participants would repeat this pre-learning task to ensure the schemas were fresh in their minds before immediately entering the MEG scanner and completing the main temporal order memory task. It took approximately 30 minutes to get participants in the scanner, resulting in a delay between the pre-learning phase and the temporal order memory task that was not present in the previous behavioural experiments. The temporal order memory task was followed by a localiser task that was used for training the classifiers.

Temporal order memory task

The temporal order memory task took the same form as that used in Experiment 4 with the ISI extended to 1000 ms. The study was split into study-test blocks with three sequences per block. Participants would study three four-item sequences that were either Congruent, Incongruent, or Unrelated to the pre-learned schemas. In contrast to the previous behavioural experiments, participants did not make animacy judgements while studying the sequences in the main task. This was to prevent motor or pre-motor responses from generating unwanted artifacts in the MEG signal. However, participants did make animacy judgements during pre-learning. Each test phase was identical to that used in Experiment 4. Participants would be cued with an exemplar from a sequence and required to recall the following two items in a four-alternative forced-choice task using buttons on a response pad that represented quadrants on the screen.

To reduce the amount of time spent in the scanner, the total number of blocks was also reduced. The task consisted of 35 blocks, eight of which were Congruent filler blocks. The remaining 27 blocks provided 27 Congruent, Incongruent, and Unrelated sequences each. This is approximately half the number of sequences in the behavioural experiments that provided 50 sequences per condition. Each of these blocks contained a single Unrelated sequence with the other two sequences either being two Congruent sequences, two Incongruent sequences, or one Congruent sequence and one Incongruent sequence. Congruent filler blocks always occurred immediately following a block containing two Incongruent sequences or on the next block but one. Multiple blocks containing two Incongruent sequences could not occur back-to-back. This was done to prevent too many incongruent sequences from occurring in a row.

For the analysis of the behavioural data, Congruent filler blocks were removed to be in line with the previous behavioural experiments. However, given that the MEG analyses attempted to detect pre-stimulus activation in the Congruent condition, the Congruent filler blocks were included in the MEG analyses to increase statistical power. This gave 27 sequences in the Incongruent and Unrelated conditions and 51 sequences in the Congruent condition. When analysing all Congruent trials (four epochs per sequence) this gave a total of 204 Congruent epochs. When analysing all Unrelated trials there were 108 epochs. When breaking the sequences down by Cue, Cue+1 and Cue+2 this gave 27 epochs each for the Incongruent and Unrelated conditions and 51 for the Congruent condition.

Localiser task

The localiser task used the exemplar discrimination task described in the previous pilot experiment. The only difference was that each of the 14 blocks now contained unique exemplars. This meant that there were a total of 56 presentations of each category, all of which were unique exemplars of that category. Which exemplars were used in the localiser (training set) and which were used in the main temporal order memory task (test set) was randomised across participants.

4.3.1.1 MEG data acquisition and processing

MEG recordings were acquired using a FieldLine HEDscan Optically-Pumped Magnetometer (OPM) system. The system had 128 sensors, however, due to sensors being down during different scanning sessions the number of active sensors varied across participants (M = 118.18, SD = 2.96, range = 111-123). Data was collected at a sampling rate of 5000 Hz. All analyses in this chapter were performed in sensor space, however structural MRIs and LIDAR scans were performed to facilitate source reconstruction at a later date. Data preprocessing and analysis were performed using the Python modules MNE Python (Gramfort, 2013), Numpy (Harris et al., 2020), and Scikit-learn (Pedregosa et al., 2011).

Raw MEG recordings were bandpass filtered between 0.5 - 45 Hz and notch filters were applied at 50 and 60 Hz to remove any residual noise from the mains frequency and projector frame rate respectively. Bad channels were rejected via visual inspection. The data were then subjected to Independent Components Analysis (ICA) and eye blink and cardiac artifacts were removed via visual inspection. Finally, data were epoched into -1000 ms to +2000 ms segments around stimulus onset for the main temporal order memory task. This created an epoch for each category that extended from the offset of the preceding category to the onset of the subsequent category. This resulted in overlapping epochs in the temporal order memory task data that could be cropped as needed for each analysis. For the localiser task, the data was epoched between -400 to +1000 ms. Data were downsampled to 250 Hz after epoching.

4.3.1.2 Time-frequency decomposition

Cue, Cue+1 and Cue+2 trials in each Congruency condition during encoding were cropped between -1000 ms to +1000 ms around stimulus onset and were baseline corrected using a time period between -900 to -750 ms. This gave 27 trials for Cue, Cue+1 and Cue+2 for the Incongruent and Unrelated condition and 51 trials for each in the Congruent condition. Each trial was then convolved with Morlet wavelets, using the MNE *compute_tfr* method, to derive the oscillatory power of the MEG signal across time and frequency. Twenty-seven frequencies ranging from 4 - 30 Hz were used, using a logarithmic spacing to distribute power across the frequency bands known to be of importance for human cognition. This logarithmic spacing ensures that frequency bands with greater widths (e.g., Beta, 13 - 30 Hz) are not prioritised over frequency bands with narrower widths (e.g., Theta, 4 - 8 Hz). For each frequency, the number of cycles was defined as half the frequency to balance time and frequency resolution across the analysed bands. The data were normalised to decibels (dB) by applying a log-ratio baseline period between -500 to –250 ms. Finally, edge artifacts were removed by cropping the epochs between -750 ms to +850 ms. The time-frequency representations were then averaged across sensors to give a grand average for each participant and condition. The raw decompositions were then contrasted in two different ways. First, Congruent > Unrelated, Incongruent > Unrelated and Incongruent > Congruent contrasts were performed for Cue, Cue+1 and Cue+2 trials respectively. Second, Cue+1 > Cue was compared within each Congruency condition. Significant clusters for each contrast were determined while controlling for multiple comparisons by using cluster-based permutation tests performing a one-sample t-test against zero. Permutation tests involved 1000 permutations.

4.3.1.3 Time-resolved decoding and time generalisation

I applied time-resolved decoding to determine whether categories in the temporal order memory task could be decoded using classifiers that were trained on the localiser data. I also ran the same analysis using a time generalisation approach to inspect how well training at one time point generalised to other time points. The localiser data was first baseline corrected between -200 to -50 ms, the temporal order memory task data was baseline corrected between -900 to -750 ms. MVPA analyses were conducted using a combination of MNE Python and Scikit-learn. All models used Linear Discriminant Analysis (LDA). Classifiers were trained for every category on the localiser task in a one-vs-all manner for each time point between -400 ms to +1000 ms around stimulus onset. There was an 87.5% class imbalance, so classifier performance was quantified using ROC AUC. Classifiers were trained on the localiser epochs with sensors as features. Prior to training, each feature was z-scored across epochs per time point. After being trained on the localiser data, classifiers were tested on the data in the temporal order memory task around epochs of -400 to +1000 ms in a timepoint-by-timepoint fashion. Classifiers were applied separately for both the Congruent and the Unrelated conditions, yielding a time series of ROC AUC metrics for each time point, category, and condition. Time-resolved decoding was applied using the

MNE *SlidingEstimator*. Time generalisation was applied using the MNE *GeneralizingEstimator*. All analyses averaged classifier performance across categories within participant. In total there were 204 trials used in the Congruent analysis and 108 trials used in the Unrelated analysis.

4.3.1.4 Pre-stimulus decoding

This analysis aimed to see if the sensor space representation present during peak decoding in the time-resolved decoding analysis could be generalised to the pre-stimulus period, providing evidence of predictive reactivation of a given category. To do this I took the localiser data and averaged the MEG signal between 150 - 350 ms post stimulus onset. This time period was selected as it is approximately ± 100 ms around the peak decoding time in the time-resolved decoding analysis. For each epoch, I extracted this single average time point to act as the training data. Classifiers were trained for each category on this single averaged time point. The localiser data was baseline corrected between -200 to -50 ms and the temporal order memory task data was baseline corrected between -900 to -750 ms prior to analysis.

In this analysis, each category classifier in the Congruent condition was trained against all other out-of-schema categories. For example, if the schemas were: (red) *face-building-dog-tool*, (blue) *furniture-car-plant-insect*, and I was training a classifier for *face*, *face* would be trained against *furniture-car-plant-insect*. As the Unrelated condition could contain any of the eight categories, no categories were strictly out-of-schema. As such, the negative class was defined as four other categories selected at random. This ensured that the negative class in both the Congruent and Unrelated condition consisted of four categories. Once trained, these classifiers were then generalised to every time point in every epoch of -1000 to +1000 ms in the temporal order memory task. For each time point a raw discriminant value was extracted. This left a data structure concerning the raw discriminant values for each category's classifier over each time point in all epochs in the temporal order memory task.

For each category classifier, I then broke the test data down based on whether the category the classifier was trained to discriminate was expected Next in the sequence, whether it was within-schema but not next, or whether it was out-of-schema. For example, imagine we were investigating the results of the *building* classifier and the schemas were: (red) *face-building-dog-tool*, (blue) *furniture-car-plant-insect.* If the *building* classifier was being applied to an epoch ± 1000 ms around a trial in which a *building* was to appear, this epoch would be labeled as *Next.* However, if the *building* classifier was applied to an epoch that occurred within the same schema as *building* but when *building* was not next, for example, on an epoch when a *dog* would appear next, that epoch would be labeled *Within-schema.* Finally, if the classifier was applied to an epoch from the opposite schema, such as when *furniture* was about to appear, that epoch would be labeled *Out-of-schema.*

This allowed me to plot the raw discriminant values from a classifier for each category depending on whether it was being applied to an epoch when that category was Next, the category was Within-schema but not next, or the category was Out-of-schema. This was done separately for both the Congruent and Unrelated sequences. Notably, in the Unrelated condition, no categories were strictly Within or Out-of-schema as the Unrelated sequences can contain any of the eight categories. Therefore, the epochs in the Unrelated analysis were either labeled as *Next* or *Not next.*

As discriminant values can differ in scale between classifiers, the discriminant values were z-scored prior to any averaging across categories and conditions. Each time series was averaged across categories within participant before being averaged across participants. An identical control analysis was also run in which the classifiers were trained on randomly labeled data.

4.3.1.5 Frequency analysis of pre-stimulus decoding time series

This analysis was based on the findings of Hetenyi et al. (2024), who found that pre-stimulus predictions manifest as increases in alpha power in a decoding time series. To investigate whether this was the case in the present experiment a similar analysis was applied to the pre-stimulus decoding time series (of z-scored discriminant values) described under the previous heading. The pre-stimulus decoding time series described above were cropped to include data between -750 ms to 0 ms. The time point -750 ms was used because this was the end of the period used for baseline correction of the data. On a trial-by-trial basis, the time series were tapered using a Hann window and transformed to the frequency domain via a Fast Fourier Transform (FFT). The power spectrum was then computed as the squared magnitude of the FFT, yielding power values in arbitrary units. As with the pre-stimulus decoding analysis above, the results in the present analysis were broken down based on whether a category was *Next*, *Within-schema*, or *Out-of-schema* in the Congruent condition. In the Unrelated condition, the results were broken down by whether a category was *Next* or *Not next*. A secondary control analysis was also run where the power spectrum was computed for decoding time series that had been derived from classifiers trained on random labels.

4.3.1.6 Cross-validation of pre-stimulus period

Although reactivation in the pre-stimulus period may be decoded using a classifier trained during stimulus perception, it is also possible that pre-stimulus activity could be category-specific, but that this activity is not representative of the patterns of activity observed during perception. Therefore, to determine whether category evidence could be decoded from the pre-stimulus period itself, a cross-validation analysis was performed on the ISIs during the main temporal order memory task. For the Congruent and Unrelated conditions, I extracted epochs between -1000 ms to 0 ms. This gave epochs containing only the pre-stimulus period, with no periods during which categories were actively perceived. For each condition two classifiers were trained, one to decode the upcoming category, and one to decode the position in the sequence (i.e., the first, second, third, or fourth ISI period in the sequence). Note, category and position were correlated in the Congruent condition, but were decoupled in the Unrelated condition, allowing me to ask whether classification in the ISI of the Congruent condition was related to the upcoming category or sequential position.

The analyses used a 5-fold cross-validation procedure with ten repetitions. A classifier was trained for each of the eight categories or four positions using a one-vs-all approach. To account for class imbalance, classifier performance was quantified using ROC AUC. Classifiers were trained on labeled MEG epochs, with sensors as features. Prior to training, each feature was z-scored across epochs. Classification was performed in a time-resolved manner by applying LDA on a timepoint-by-timepoint basis, yielding an ROC AUC metric for each time point and category / position. The resulting time series were then averaged across categories / positions within participant to derive a single time series for each participant.

4.3.2 Results

4.3.2.1 Behavioural results

4.3.2.1.1 Pre-learning

A trials-to-criterion analysis was run on the pre-learning data to establish whether participants were reaching criterion faster in session 1 than in session 2. Many participants were close to ceiling in the first session, resulting in a violation of normality of differences. As such, median and interquartile range values are reported with a Wilcoxon test. There was no significant difference between the median number of trials to reach criterion in session

1 (*Median* = 22, *IQR* = 1.75) and session 2 (*Median* = 22, *IQR* = 0), V = 100.5, p = .098, r = 0.32. This indicates that participants may not have been retaining information from session 1. However, the large number of participants near ceiling in session 1 (n = 26) suggests that this lack of a difference is likely due to the ease of the task, rather than a lack of memory between sessions.

4.3.2.1.2 Retrieval

Mean accuracy across the Congruency and Test manipulations at retrieval can be seen in Table 4.1. A full summary of the mixed-effects model at retrieval can be seen in Table 4.2. The results of the retrieval analysis are plotted in Figure 4.4.

Table 4.1: Mean percentage accuracy at retrieval between Cue+1 and Cue+2 across the three Congruency conditions at the category, exemplar, and conditional exemplar level in the MEG experiment. Standard deviations are shown in parentheses.

Level	Congruency	Cue+1	Cue+2
Category	Unrelated	65.33 (14.09)	68.21 (10.82)
	Congruent	83.95 (12.07)	80.25 (11.2)
	Incongruent	72.63 (11.95)	62.96 (13.77)
Exemplar	Unrelated	51.75 (17.94)	52.88 (15.78)
	Congruent	62.86 (17.45)	59.57 (16.38)
	Incongruent	58.74 (16.87)	49.18 (17.87)
Conditional Exemplar	Unrelated	80.28 (11.73)	77.98 (14.09)
	Congruent	74.8 (12.85)	75.08 (11.54)
	Incongruent	81.4 (11.91)	78.01 (15.82)

Category level

To be included in the category-level retrieval analysis participants were required to be significantly above chance (50%) in terms of category-level accuracy collapsed across Congruency and Test manipulations. This resulted in a final sample size of n = 36.

When collapsing across Test, accuracy was significantly higher on Congruent trials as compared to both the Unrelated (β = 0.86, *SE* = 0.08, z = 11.06, p <.001) and Incongruent conditions (β = 0.8, *SE* = 0.08, z = 10.27, p <.001). However, there was no difference between the Incongruent and Unrelated conditions (β = 0.06, *SE* = 0.07, z = 0.81, p = .417).

However, there were two significant Congruency by Test interactions. Test interacted with both the Congruent vs. Unrelated contrast ($\beta = -0.4$, *SE* = 0.16, z = -2.55, p = .011) and the Incongruent vs. Unrelated contrast ($\beta = -0.6$, *SE* = 0.14, z = -4.29, p <.001).

When assessing the effects of Congruency at Cue+1 alone, all three contrasts were significant. Accuracy in the Congruent condition was higher than in the Unrelated ($\beta = 1.06$, *SE* = 0.11, z = 9.45, p <.001) and Incongruent conditions ($\beta = 0.7$, *SE* = 0.11, z = 6.11, p <.001). Crucially, accuracy was higher in the Incongruent condition relative to the Unrelated baseline ($\beta = 0.36$, *SE* = 0.1, z = 3.56, p <.001). This indicates that the MEG experiment successfully replicated the incongruency advantage observed in Experiment 4.

When assessing the effects at Cue+2, the Congruency effect was still present, with accuracy in the Congruent condition being higher than in the Unrelated ($\beta = 0.66$, *SE* = 0.11, z = 6.15, p <.001) and Incongruent condition ($\beta = 0.91$, *SE* = 0.11, z = 8.53, p <.001). However, at Cue+2 the effect of incongruency was flipped, and there was a significant Incongruency disadvantage when comparing the Incongruent and Unrelated conditions ($\beta = -0.24$, *SE* = 0.1, z = -2.49, p = .013).

Exemplar level

To be included in the exemplar-level retrieval analysis participants were required to be significantly above chance (25%) in terms of exemplar-level accuracy collapsed across Congruency and Test manipulations. This resulted in a final sample size of $n = 36$.

When collapsing across Test, the effects at the exemplar level mirrored those observed at the category level. Accuracy was higher in the Congruent condition relative to both the Unrelated ($\beta = 0.4$, *SE* = 0.07, z = 5.84, p <.001) and Incongruent conditions ($\beta = 0.32$, *SE* = 0.07, z = 4.75, p <.001). There was no difference between the Incongruent and Unrelated conditions ($\beta = 0.07$, *SE* = 0.07, z = 1.09, p = .275).

I also observed a Congruency by Test interaction for the Incongruent vs. Unrelated contrast ($\beta = -0.47$, *SE* = 0.13, z = -3.5, p <.001). However, there was no interaction for the

Congruent vs. Unrelated contrast at the exemplar level ($\beta = -0.2$, $SE = 0.14$, z = -1.47, $p = .140$).

When breaking the model down by Test, I saw a similar pattern of results at Cue+1 as were observed at the category level. Accuracy was higher in the Congruent ($\beta = 0.5$, $SE = 0.1$, $z = 5.16$, $p < .001$) and Incongruent conditions ($\beta = 0.31$, $SE = 0.1$, $z = 3.24$, $p = .002$) relative to the Unrelated baseline. However, although accuracy was numerically higher in the Congruent condition relative to the Incongruent condition, this difference did not reach statistical significance ($\beta = 0.19$, $SE = 0.1$, $z = 1.94$, $p = .053$).

At Cue+2, accuracy was higher in the Congruent condition relative to both the Unrelated ($\beta = 0.3$, $SE = 0.1$, $z = 3.1$, $p = .004$) and Incongruent conditions ($\beta = 0.46$, $SE = 0.1$, $z = 4.8$, $p < .001$). However, there was no difference between the Incongruent and Unrelated conditions ($\beta = -0.16$, $SE = 0.09$, z = -1.71, $p = .087$).

Conditional exemplar level

To be included in the conditional exemplar-level analysis participants were required to be above chance performance at the exemplar level when only considering trials on which category-level information was accurately recalled (50%). This resulted in a final sample size of $n = 33$.

Interestingly, in the MEG experiment I observed a Congruency disadvantage at the exemplar level. When collapsing across Test, accuracy was significantly lower in the Congruent condition relative to both the Unrelated ($\beta = -0.25$, $SE = 0.1$, z = -2.56, $p = .021$) and the Incongruent conditions ($\beta = -0.29$, $SE = 0.1$, z = -3.02, $p = .008$). However, there was no difference between the Incongruent and Unrelated condition ($\beta = 0.05$, $SE = 0.11$, $z = 0.44$, $p = .659$).

There were no interactions in the conditional exemplar model, however, I broke the model down by Test for completeness. At Cue+1, accuracy was lower in the Congruent condition relative to the Unrelated ($\beta = -0.33$, $SE = 0.14$, z = -2.35, $p = .038$) and the Incongruent conditions ($\beta = -0.41$, $SE = 0.14$, z = -2.99, $p = .008$). However, there was no difference between the Incongruent and the Unrelated condition ($\beta = 0.08$, $SE = 0.15$, $z = 0.54$, $p = .588$). This indicates that there was a significant congruency disadvantage for exemplar-level information at Cue+1. Notably, this finding is in line with SLIMM's prediction of reduced memory for detail under schema-congruent conditions. I discuss this point in more detail in the discussion. There were no significant contrasts at Cue+2 in the conditional exemplar model.

Table 4.2: Summary of the mixed-effects model at retrieval for the category, exemplar, and conditional exemplar-level analyses in the MEG experiment. P-values are uncorrected and therefore may differ from those in the text.

Level	Fixed effect	Estimate	SE	z	p
Category	Intercept	0.68	0.11	6.19	<.001
	Congruent (vs. unrelated)	1.06	0.11	9.45	<.001
	Incongruent (vs. unrelated)	0.36	0.10	3.56	<.001
	Test (Cue+1 vs. Cue+2)	0.14	0.10	1.38	.167
	Congruent x Test	-0.40	0.16	-2.55	.011
	Incongruent x Test	-0.60	0.14	-4.29	<.001
Exemplar	Intercept	0.09	0.13	0.70	.484
	Congruent (vs. unrelated)	0.50	0.10	5.16	<.001
	Incongruent (vs. unrelated)	0.31	0.10	3.24	.001
	Test (Cue+1 vs. Cue+2)	0.05	0.09	0.52	.601
	Congruent x Test	-0.20	0.14	-1.47	.140
	Incongruent x Test	-0.47	0.13	-3.50	<.001
Conditional Exemplar	Intercept	1.51	0.15	10.14	<.001
	Congruent (vs. unrelated)	-0.33	0.14	-2.35	.019
	Incongruent (vs. unrelated)	0.08	0.15	0.54	.588
	Test (Cue+1 vs. Cue+2)	-0.14	0.15	-0.98	.330
	Congruent x Test	0.16	0.19	0.80	.421
	Incongruent x Test	-0.07	0.21	-0.33	.739

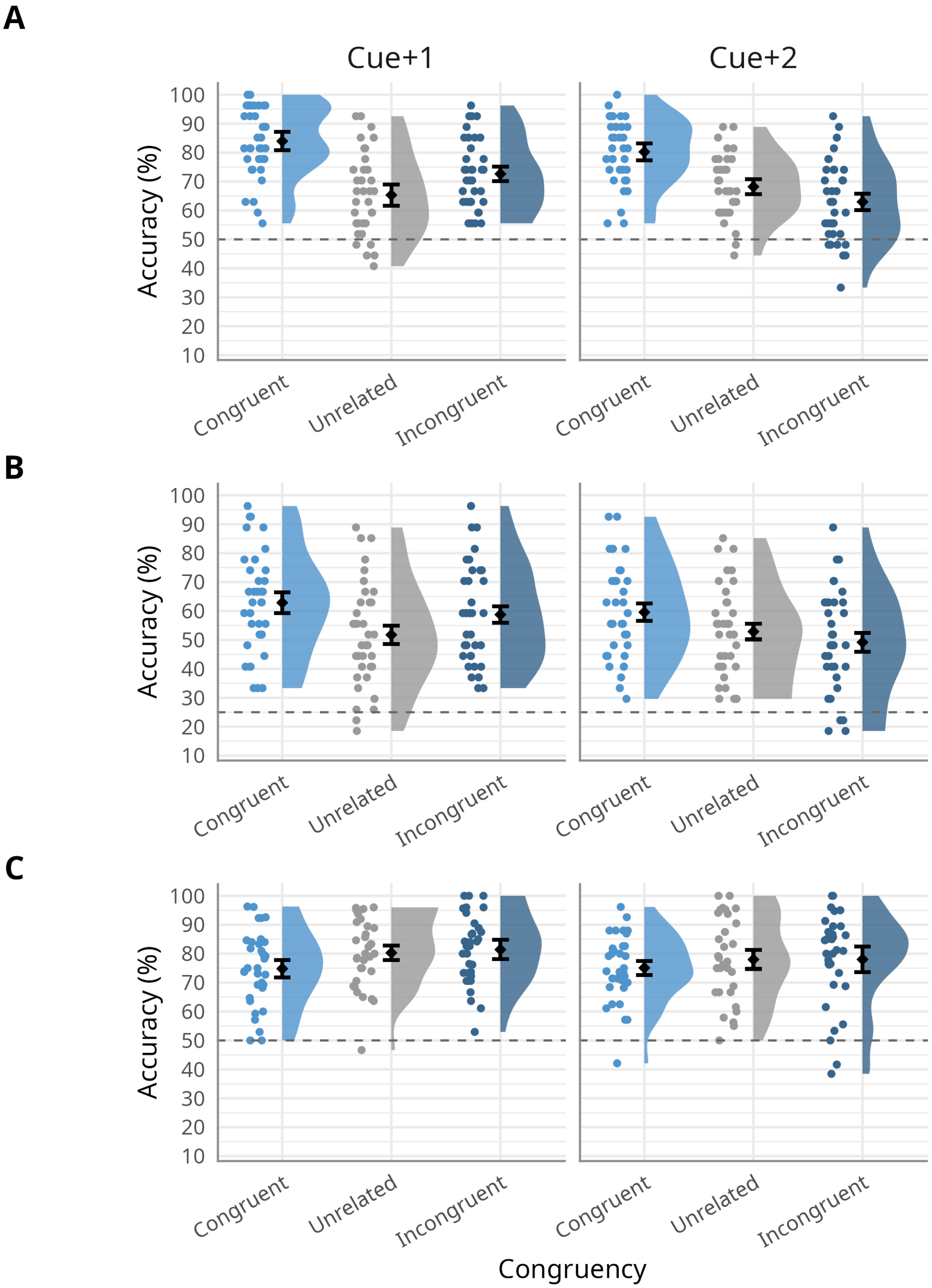

Figure 4.4: Mean accuracy (%) at retrieval across the Congruency and Test manipulations at (A) category level, (B) exemplar level, and (C) conditional exemplar level in the MEG experiment. Diamonds show the sample means. Density and scatter plots show the distribution of participant means. Error bars represent 95% confidence intervals.

4.3.2.2 MEG Results

Four participants were excluded from the MEG analyses. One participant was excluded due to a software error where their data from the localiser task failed to save. One participant was excluded for having excessively noisy data due to movement artifacts. One participant asked for a scanning run to be terminated and one participant was excluded for excessive head movement. This resulted in a final sample size of $n = 33$ for the MEG analyses.

4.3.2.2.1 Time-frequency

First, I aimed to investigate whether there was any evidence of prediction error signals in the time-frequency representations. A time-frequency representation for each Congruency condition at Cue, Cue+1 and Cue+2 respectively can be found in Figure 4.5

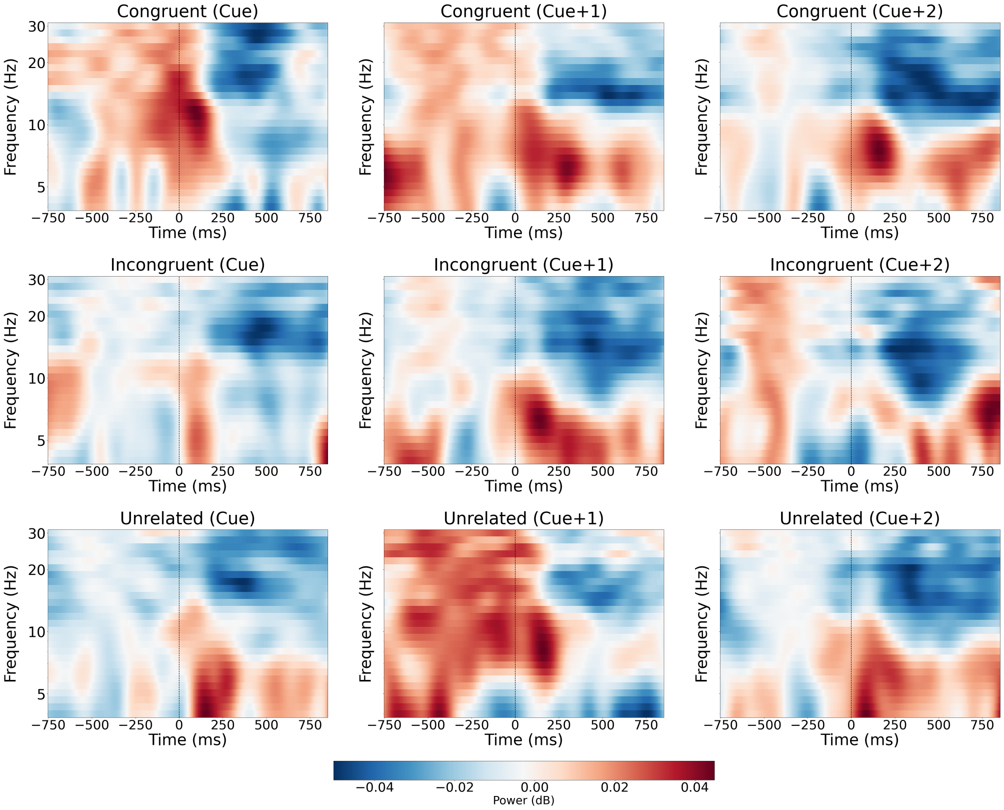

Figure 4.5: Time-frequency plots showing the Congruent (top row), Incongruent (middle row), and Unrelated conditions (bottom row) at Cue (left column), Cue+1 (middle column), and Cue+2 (right column) from -750 to +800 ms. Dotted vertical line shows stimulus onset. Power is displayed in decibels. Freqeuncies on y-axis are log-spaced.

I contrasted the time-frequency representations (TFR) for the Congruent, Incongruent, and Unrelated conditions in a pairwise manner at Cue, Cue+1, and Cue+2 respectively to see how power differed between the conditions at each position in the sequence. The results from this analysis are plotted in Figure 4.6.

Each TFR contrast was subjected to a cluster-based permutation test that performed a one-sample t-test against zero. A single significant cluster was found. At Cue+1 in the Incongruent condition, the trial on which the violation occurred, there was an increase in theta power (4 - 8 Hz) from approximately 250 ms to 800 ms. This finding is consistent with research that suggests that theta oscillations signal prediction errors. Notably, this increase in theta power in the Incongruent condition is not present at Cue+2, despite the fact that Cue+2 was also schema incongruent. This aligns well with the behavioural data, where there was an incongruency advantage at Cue+1 but not at Cue+2. However, it is noteworthy that there were no significant clusters when contrasting the Incongruent condition with the Congruent condition, despite theta appearing numerically higher in the Incongruent condition at Cue+1.

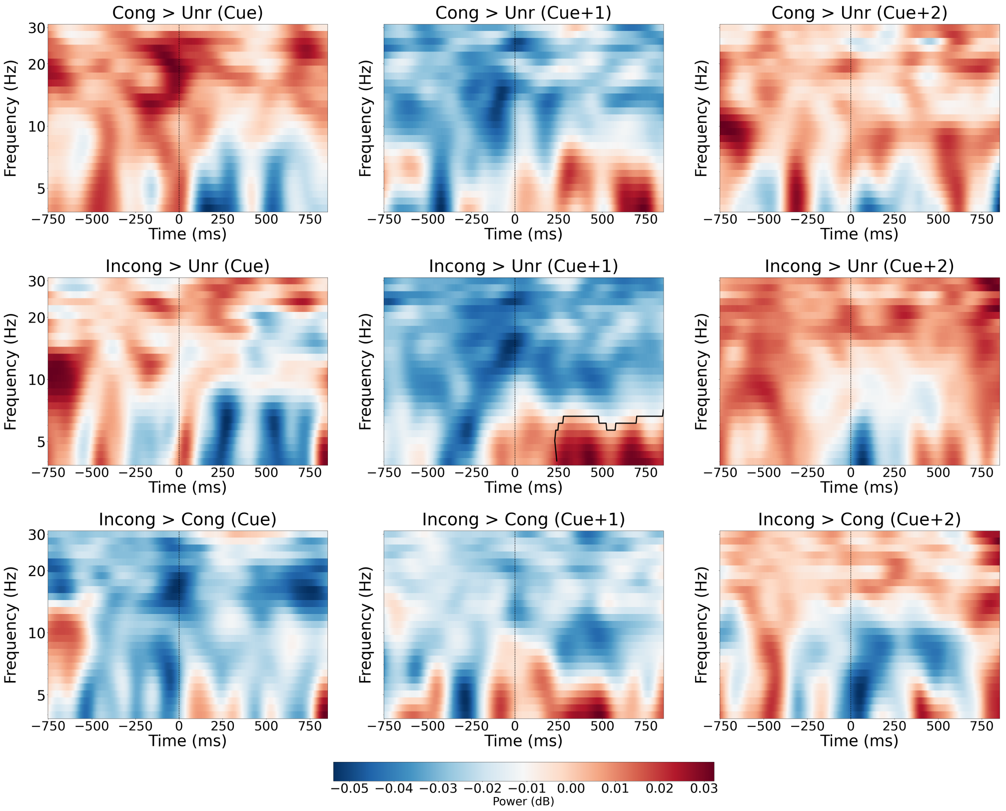

Figure 4.6: Time-frequency plots showing the contrasts Congruent > Unrelated (top row), Incongruent > Unrelated (middle row), and Incongruent > Congruent (bottom row) at Cue (Left column), Cue+1 (middle column), and Cue+2 (right column) from -750 to +800 ms. Dotted vertical line shows stimulus onset. Black contours show significant clusters as determined by cluster-based permutation tests performing a one-sample two-tailed t-test against zero. Power is displayed in decibels. Freqeuncies on y-axis are log-spaced.

Finally, I investigated whether there were differences between Cue and Cue+1 in each of the three Congruency conditions. The results of this analysis are plotted in Figure 4.7. There was a single significant cluster in this analysis in the Unrelated condition, in which there was a significant decrease in theta-band activity at Cue+1. This finding suggests that the theta signal observed in the Incongruent > Unrelated contrast in the previous analysis may have, at least in part, been driven by a relative decrease in theta between Cue and

Cue+1 in the Unrelated condition, rather than an increase in theta in the Incongruent condition per se. Notably, this pattern is still in line with the U-shaped function observed in the behavioural data. Research suggests that higher theta may be associated with greater encoding (Joensen et al., 2023; Kota et al., 2020; Lin et al., 2017). A significant decrease in theta in the Unrelated condition at Cue+1 may explain why memory retention is impaired in the Unrelated case. It is not entirely clear why theta would decrease in the Unrelated condition at Cue+1, however, one possibility is that it is a sequential effect. Cue+1 is always the second or third item in the sequence, it is never the first or last item, and the decrease in theta may represent less attention toward central items. I return to this point in the discussion.

The time-frequency analyses suggest that differences in theta-band activity are present during encoding and are in line with the behavioural U-shaped function of schema congruency. However, these effects seem to be partially driven by decreases in theta in the Unrelated condition, which calls into question whether the significant cluster in the Incongruent > Unrelated contrast truly reflects a prediction error signal. However, it is worth highlighting that this analysis involved averaging time-frequency representations across sensors to derive a sensor space grand average. This averaging may reduce power and future analyses using topographies or source space data may provide stronger evidence for differences.

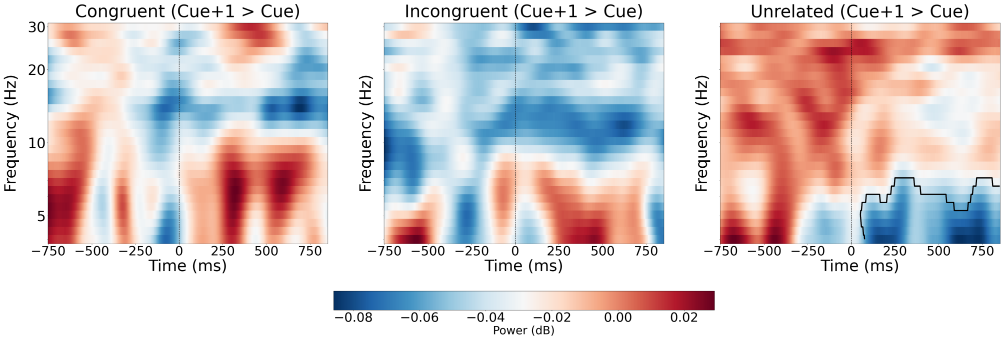

Figure 4.7: Time-frequency plots showing the contrasts Cue+1 > Cue in the Congruent (Left), Incongruent (middle), and Unrelated (right) conditions from -750 to +800 ms. Dotted vertical line shows stimulus onset. Black contours show significant clusters as determined by cluster-based permutation tests performing a one-sample two-tailed t-test against zero. Power is displayed in decibels. Freqeuncies on y-axis are log-spaced.

4.3.2.2.2 Time-resolved decoding

This analysis was designed to demonstrate that the categories were indeed decodable using a train-test decoding paradigm (rather than the cross-validation paradigm used in the pilot). In doing so, I could determine the peak decoding time points that would be used to train classifiers in the pre-stimulus decoding analyses. Results from the time-resolved decoding analysis are plotted for the Congruent and Unrelated conditions in Figure 4.8.

Consistent with prior research (T. Carlson et al., 2013) category evidence began to steeply rise within the first 100 ms after stimulus onset. This was true for both the Congruent and Unrelated condition. When averaging the decoding time series timepoint-by-timepoint peak decoding accuracy was similar in both the Congruent ($M = 0.59$) and Unrelated conditions ($M = 0.59$). Likewise peak decoding occurred at similar times in each condition. In the Congruent condition peak decoding occurred at 248 ms, whereas peak decoding in the Unrelated condition occurred at 236 ms.

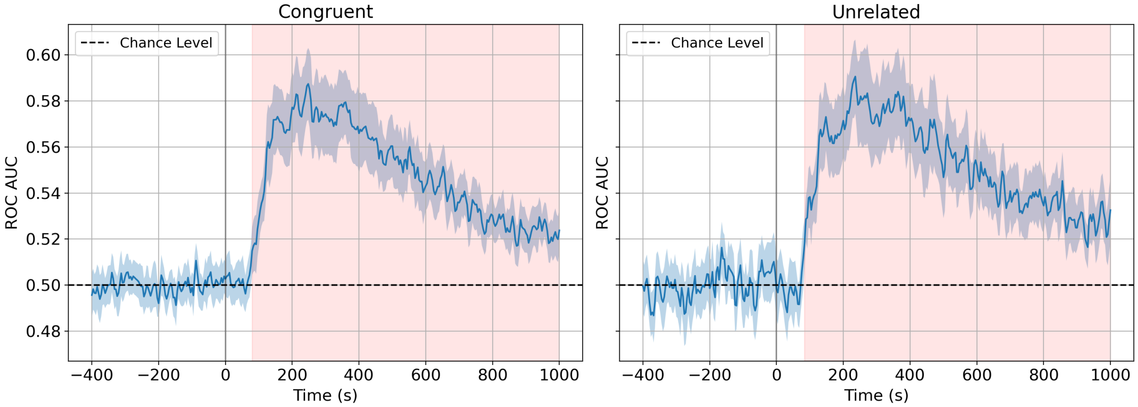

Figure 4.8: Time series showing the time-resolved decoding averaged across category when training on the localiser data and testing on the data from the main temporal order memory task in the Congruent condition (Left) and the Unrelated condition (Right). Vertical line indicates stimulus onset, Horizontal line indicates chance level performance. Blue shaded areas show 95% confidence intervals. Red shaded area shows time points of significant decoding as determined by cluster-based permutation tests.

Notably, averaging timepoint-by-timepoint across participants may reduce classifier fidelity, as different participants may peak at different times. To address this, I ran an analysis looking at the peak decoding value for each participant regardless of when that occurred. Overall, peak decoding was numerically lower in the Congruent condition (M = 0.62, SD = 0.04) than in the Unrelated condition (M = 0.65, SD = 0.04), a difference that was statistically significant $t(32)$ = -5.93, p <.001, d = -0.58. These results are plotted in Figure 4.9.

Why average decoding peaks would be higher in the Unrelated condition is not entirely clear. However, if predictions of future categories are taking place in the Congruent condition, this may weaken the representation of the current category somewhat. I return to this point in the discussion.

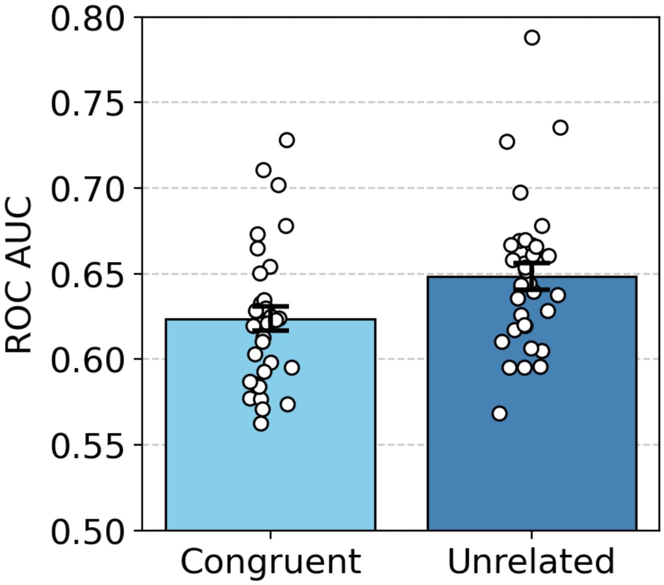

Figure 4.9: Bar chart showing the average peak decoding value in the time-resolved decoding analysis in the Congruent and Unrelated conditions. Dots represent individual participant means. Error bars represent 95% confidence intervals.

4.3.2.2.3 Time generalisation

Next, I used time generalisation to investigate how classifiers trained at specific time points generalised to other times. Notably, this would also allow for the first exploration of pre-stimulus decoding as classifiers trained on time points post stimulus onset would be tested on timepoints in the pre-stimulus period. For every category and timepoint in a -400 to 1000 ms time window, a classifier was trained on the localiser data and this classifier was tested on every time point in that time frame in the temporal order memory task. This resulted in a time generalisation matrix for each category which were then averaged across categories to give a final time generalisation matrix for each participant. Results from this time generalisation analysis are plotted in Figure 4.10.

Cluster-based permutation tests performing a two-tailed one-sample t-test against baseline (ROC AUC = 0.5) were applied to assess clusters of significant above/below chance classification in both the Congruent and Unrelated conditions. In both analyses there were two significant clusters. The first cluster, prominent in both conditions, revealed the typical pattern observed in a time generalisation analysis (King & Dehaene, 2014) in which classifier performance peaked when classifiers were tested on similar timepoints to which they were trained. Notably, the cluster encompassed most of the train and test times after approximately 100 ms, indicating that classifiers did generalise across time, suggesting a relatively stable mental representation.

Of particular note was a second cluster that indicated decoding in the pre-stimulus period. Interestingly, ROC AUC seemed to be significantly below baseline as stimulus onset approached, suggesting that classifiers were actively predicting the negative class. Although somewhat counter intuitive, a pattern that negatively correlates with an expected pattern is still predictive of that pattern, suggesting that this may be evidence of a predictive signal. However, a cluster of decoding was also present in the pre-stimulus period of the Unrelated condition, where predictions regarding the upcoming category should not be possible. As such, it is difficult to say whether this pattern is a true predictive signal or an artifact in the data.

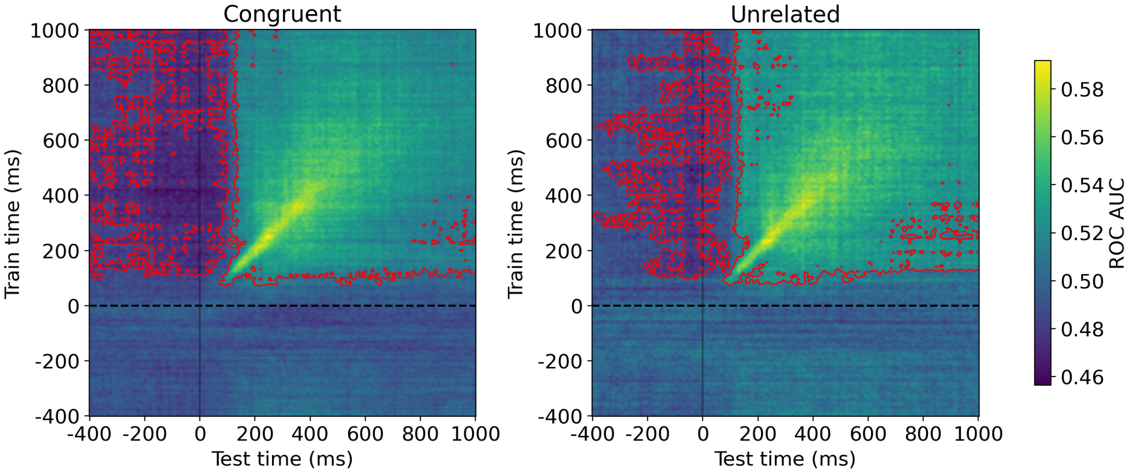

Figure 4.10: Time generalisation matrices showing category decoding averaged across categories obtained by training and testing classifiers on all pairwise time points. Red contours indicate significant clusters as determined by cluster-based permutation test performing a one-sample two-tailed t-test against baseline performance (ROC AUC = 0.5). Dashed lines indicate stimulus onset.

4.3.2.2.4 Pre-stimulus decoding

The time generalisation analysis presented evidence of pre-stimulus decoding. However, its presence in both the Congruent and Unrelated conditions, and the fact that the signal was significantly below baseline (active prediction of the negative class), made it unclear as to whether this was a true predictive signal. Therefore, I moved on to try and refine the pre-stimulus decoding analyses further. One way to determine whether the pre-stimulus activity is indeed category-specific, is to see how well a classifier performs when applied to an epoch when a different category is next. If the below chance decoding observed above is indeed category-specific, it should only be observed when a category is expected to occur next and not when it is applied to epochs relating to other categories.

Importantly, this analysis would also help to provide evidence as to whether predictions were occurring multiple steps into the future. If predictions were occurring multiple steps into the future, as the behavioural experiments suggest, this may manifest as a period of significant pre-stimulus decoding for within-schema categories that are expected in the current context but that are not next in the sequence. As such, one possible pattern of results would be strong pre-stimulus decoding for the category which is next, relatively weaker decoding of within-schema categories, and no indication of decoding for out-of-schema categories.

To do this, I trained a single classifier on peak classifier performance in the localiser task and generalised that classifier to an entire two second epoch -1000 to +1000 ms around stimulus onset in the temporal order memory task. This classifier was trained on a single vector of sensor space activity that was derived by averaging the data in each training epoch between 150 to 350 ms after stimulus onset. This time period was approximately ± 100 ms around peak classifier performance in the time-resolved decoding analysis. This averaging was aimed at reducing category-irrelevant noise that may occur across MEG samples, to produce a more veridical category-specific representation.

A classifier was trained for each category and then applied to every epoch in the main temporal order memory task and time series of raw discriminant values were retrieved. There were therefore eight time series for each epoch, one for each category-specific classifier. In the Congruent condition, for each epoch, the classifiers could then be sorted based on whether the classifier was the Next category in the sequence (i.e., was the correct label for this epoch), was a within-schema category but was not next, or was an out-of-schema category. In the Unrelated condition, the time series were sorted based on whether a category was Next or Not next. An identical procedure was also run with classifiers that were trained on random labels. Periods of significant decoding were determined using cluster-based permutation test performing a one-sample two-tailed t-test against baseline ($z = 0$). Results from the pre-stimulus decoding analysis are plotted in figure Figure 4.11.

I found that the z-scored discriminant values peaked at approximately 250 ms post stimulus onset for the time series regarding categories that were Next. This peak was also not present when the classifiers were trained on random labels providing further evidence that this effect is indeed category-specific. Once again, there was a significant negative effect away from average performance in the pre-stimulus period when a category was Next. However, as with the time generalisation analysis, this appeared in both the Congruent and Unrelated condition, raising doubt as to the precise nature of this effect.

Notably, the Within-schema and Out-of-schema time series varied little from the average discriminant score in the pre-stimulus period, indicating that Within and Out-of-schema categories could not be decoded during the pre-stimulus period. Moreover, the negative effect was not present for classifiers trained on random labels. These findings suggest that the negative effect in the pre-stimulus period is indeed category-specific, and not just an artifact observed for any classifier applied to the pre-stimulus period. However, the precise nature of this effect and why it appears in both the Congruent and Unrelated condition remains unclear. Additionally, the lack of decoding for Within-schema categories means this analysis has provided little evidence for predictions that occur multiple steps into the future.

Finally, it is worth highlighting that there were significant below average discriminant values post-stimulus for Out-of-schema categories, an effect that was not present for Within-schema categories. This effect is likely due to the fact that the Out-of-schema categories were included in the negative class but the Within-schema categories were not.

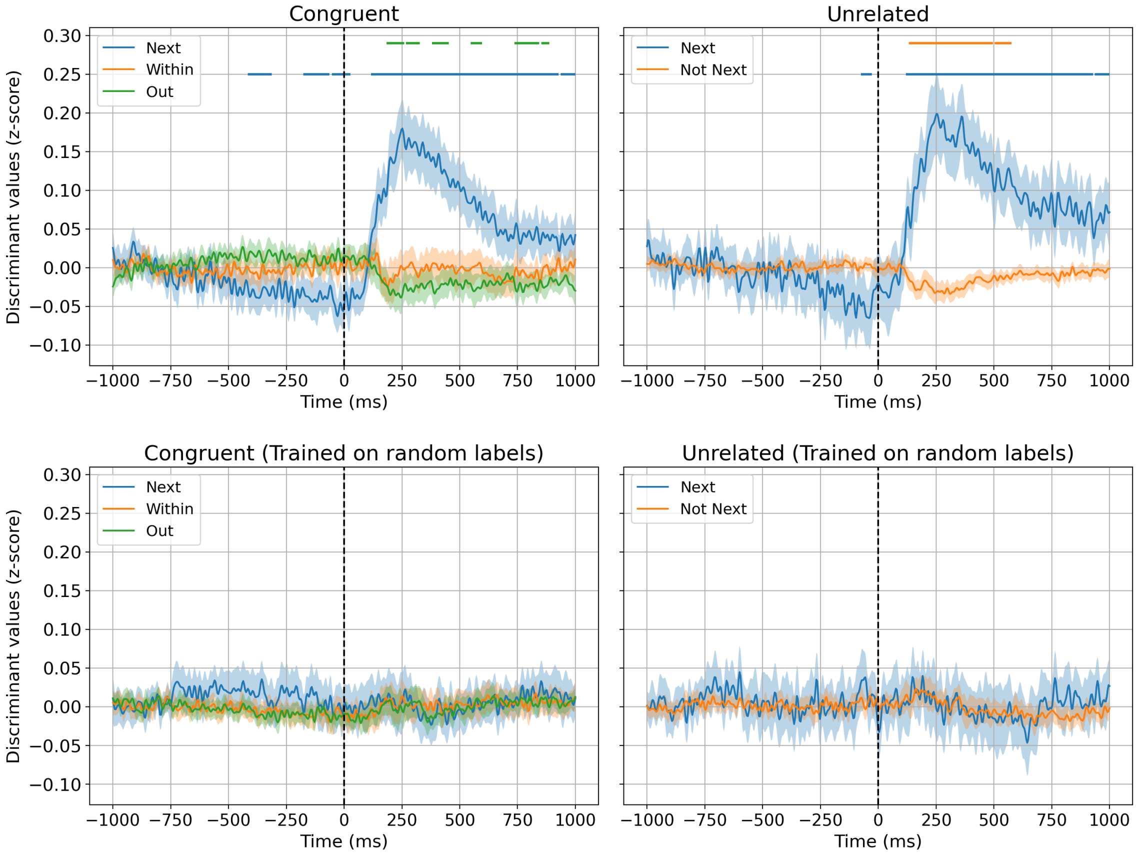

Figure 4.11: Decoding time series averaged across category showing the z-scored discriminant values. Classifiers were trained for each category and then applied to every epoch in the main temporal order memory task. In the Congruent condition (Left) these epochs were broken down by whether a category-specific classifier was being applied to an epoch when the category was Next (blue), when the category was Within-schema but not next (orange), or whether the category was Out-of-schema (not expected in the current sequence) (green). In the Unrelated condition (Right), the epochs were broken down by whether a category-specific classifier was being applied to an epoch when the category was Next (blue) or when it was Not Next (orange). Dashed vertical line indicates stimulus onset. Shaded areas show 95% confidence intervals. Horizontal lines above the time series indicate periods of significant decoding as determined by cluster-based permutation test.

4.3.2.2.5 Frequency analysis of pre-stimulus decoding time series

Hetenyi et al. (2024) found that pre-stimulus predictions manifested not in overall classifier performance but in the frequency characteristics of the decoding time series. Specifically, they found increased alpha power in a decoding time series for a predictable shape category. I therefore investigated whether this was true of the pre-stimulus time series presented in the previous analysis. To do this, I took the time series presented above in the pre-stimulus decoding section and cropped the time series to only include the period between -750 to 0 ms, the pre-stimulus period that was not used for baseline correction. I then applied an FFT to these time-series on a trial-by-trial basis. Power values were derived for each frequency between 2 - 30 Hz as the squared magnitude of the FFT values. An identical procedure was also applied to the time series trained on randomly labeled data. Results from the FFT analysis of the pre-stimulus decoding time series are plotted in Figure 4.12.

To determine whether there were any differences in alpha power across Congruency conditions and whether the classifiers were trained on accurately or randomly labelled data, power was averaged across 8 - 12 Hz for each participant and a 2x2 repeated-measures ANOVA with Congruency (Congruent vs. Unrelated) and Training labels (Accurate vs. Random) as predictors was conducted. There was a main effect of Training labels $F(1, 32) = 7.94$, $p = .008$, $\eta^2 < .01$, but no main effect of Congruency $F(1, 32) = 0.01$, $p = .928$, $\eta^2 < .01$. The Congruency by Training labels interaction failed to reach significance $F(1, 32) = 3.32$, $p = .078$, $\eta^2 < .01$.

For completeness, t-tests were conducted to determine differences in alpha power between Accurate and Random labels within each Congruency condition. In the Congruent condition alpha power for accurate labels ($M = 0.77$, $SD = 0.26$) did not differ from alpha power for random labels ($M = 0.76$, $SD = 0.23$), $t(32) = 1.02$, $p = .315$, $d = 0.07$. However, in the Unrelated condition, alpha power for accurate labels ($M = 0.81$, $SD = 0.31$) was higher than for random labels ($M = 0.73$, $SD = 0.24$), $t(32) = 2.82$, $p = .008$, $d = 0.26$.

These results indicate that increases in alpha power were only observed in the Unrelated condition. This conflicts with previous findings that have suggested this alpha signal is a marker of pre-stimulus predictions, as I only observed the effect in the Unrelated condition where upcoming stimuli are unpredictable. A potential account for this finding is reviewed in the discussion.

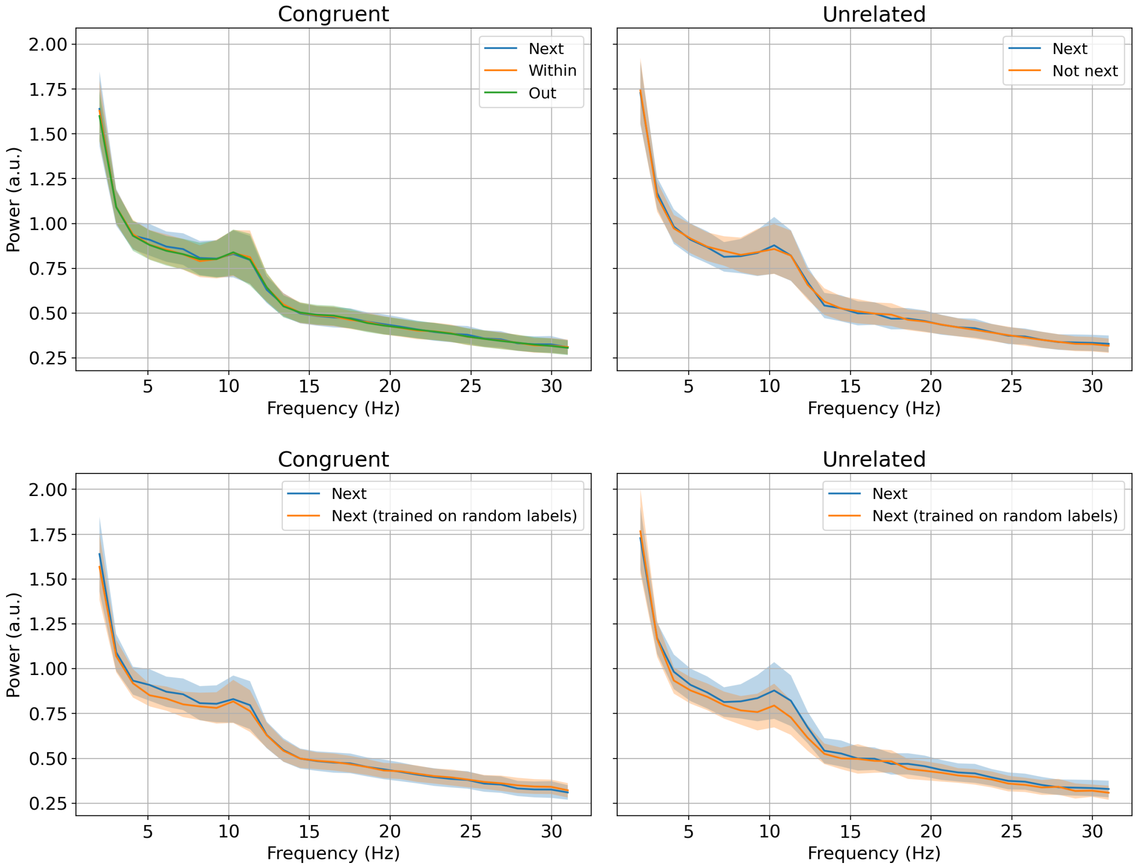

Figure 4.12: Power spectrum of the pre-stimulus decoding time series at -750 ms to 0 ms (stimulus onset). The top row shows the Congruent condition (Left) and the Unrelated condition (Right). In the Congruent condition (Left) these epochs are broken down by whether a category-specific classifier was being applied to an epoch when the category was Next (blue), when the category was Within-schema but not next (orange), or when the category was Out-of-schema (not expected in the current sequence) (green). In the Unrelated condition (Right), the epochs are broken down by whether a category-specific classifier was being applied to an epoch when the category was Next (blue) or when it was Not Next (orange). The bottom row shows the same power spectrum for time series when a category was Next, plotted against a classifier that was trained on randomised labels. Shaded areas show 95% confidence intervals.

4.3.2.2.6 Cross-validation of pre-stimulus period

Finally, it is possible that predictive signals occurring in the pre-stimulus period may not be representative of the patterns of activity that are present during stimulus perception. As such, I ran an analysis investigating whether there was any information in the pre-stimulus period itself that was predictive of the upcoming category, but that could not be decoded using a classifier trained while the category was actively perceived. To do this, I trained and tested classifiers using 5-fold cross validation on the 1000 ms pre-stimulus period prior to a category's occurrence in the main temporal order memory task. As this analysis involved cross-validation on the pre-stimulus time series, no baseline correction was applied.

I ran two analyses, one in which the data were labeled in terms of the upcoming category, and one in which the data were labeled in terms of an epoch's position in the sequence (i.e., whether it was the first, second, third, or fourth epoch in a sequence). Labeling the data in these two ways is important because in the Congruent condition the upcoming category correlates with sequence position. For example, in a Congruent sequence of the form: *face-building-dog-car, face* is always first, *building* is always second and so on. However, this is not the case in the Unrelated condition, where categories can theoretically appear in any position in the sequence. As such, it is important to label the data both in terms of the upcoming category, and in terms of sequence position, in order to disentangle what information is actually being decoded. Results from the cross-validation analysis are plotted in Figure 4.13.

When labeling the data in terms of the upcoming category I observed above chance decoding accuracy, especially in the first half of the pre-stimulus period in the Congruent condition. Notably, this was not the case in the Unrelated condition in which classifier performance was at chance throughout the pre-stimulus period. However, when labeling the data in terms of sequence position, I observed above chance decoding in both the Congruent and

Unrelated conditions, particularly in the first half of the pre-stimulus period. This pattern of results indicates that information is indeed decodable during the pre-stimulus period using a cross-validation approach. However, this decoding evidence is most likely driven by the epoch's position in the sequence rather than the upcoming stimulus category.

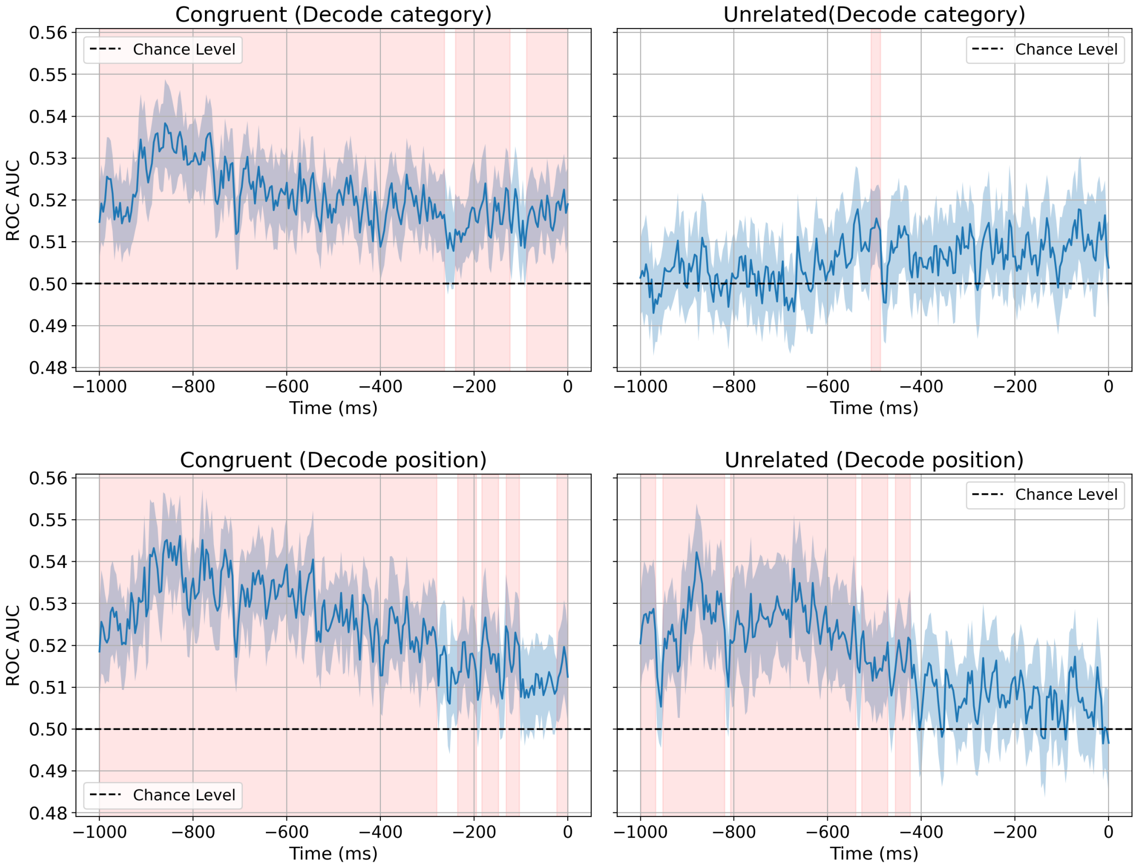

Figure 4.13: Time series showing the cross-validation of the pre-stimulus period from -1000 to 0 ms (stimulus onset) in the Congruent (Left) and Unrelated (Right) conditions. The top row shows results when the data were labeled in terms of the upcoming category. The bottom row shows results when the data were labeled in terms of position in the sequence (i.e., first, second, third, or fourth epoch in the sequence). Horizontal line shows chance level performance. Blue shaded areas show 95% confidence intervals. Red shaded areas show periods of significant decoding as determined by cluster-based permutation tests.

4.3.3 Discussion

The work presented in this chapter was aimed at replicating the U-shaped function of schema congruency observed in Experiment 4, providing evidence of prediction error signals on schema-incongruent trials, and investigating whether these effects are preceded by pre-stimulus reactivation of predictable categories.

The behavioural results successfully replicated those observed in Experiment 4. At the category level, a U-shaped function of schema congruency was observed, with higher recall accuracy in the Congruent and Incongruent conditions relative to the Unrelated baseline. This finding helps to confirm that categorical violations in particular, as opposed to sequential violations, enhance memory retention. I will return to the implications of this finding in the general discussion chapter.

Of particular note in the present study, was the finding of a significant disadvantage in the Congruent condition at the conditional exemplar level relative to both the Unrelated and Incongruent Conditions at Cue+1. This finding is of note as it is in line with a prediction of SLIMM (van Kesteren et al., 2012). SLIMM predicts that under schema-congruent conditions memory encoding should be gist-like due to mPFC-dominated encoding that bypasses hippocampal encoding of episodic detail. Likewise, under schema-incongruent conditions, hippocampal encoding should facilitate detailed memory traces. It is possible that the congruency disadvantage observed here is reflective of this gist-like vs. detailed encoding. However, this effect was not been observed in any of the previous experiments, including Experiment 4, which makes the effect difficult to dissociate from type I error.

There are two key differences that may have contributed to enhanced encoding of detail in the MEG experiment. First, there were no animacy judgements required at encoding, which may have allowed more focus on the exemplars themselves rather than generating responses to animacy-level information. Second, the MEG experiment only had 35 blocks, which was markedly less than the 64 blocks in Experiment 4. It may be the case that the additional cognitive fatigue involved in the longer experiment resulted in less engagement from participants, who fell back on cognitively easier strategies such as focusing on category-level, rather than exemplar-level, information. If participants were indeed engaging with exemplar-level information more in the MEG experiment, SLIMM would predict that this should primarily manifest as increases in performance in the Incongruent condition. In line with this, the congruency disadvantage in the present experiment seems to be driven by an increase in performance in the Incongruent condition between the two experiments, rather than a decrease in performance in the congruent condition, indicating more detailed encoding in the Incongruent condition specifically. Although somewhat speculative, it may be beneficial for future research to repeat this experiment with no animacy judgements during encoding and with the decreased number of blocks to see whether the decrease in cognitive load would result in a replication of the difference in memory for detail in the Congruent and Incongruent conditions.

Time-frequency analyses revealed two significant clusters. First, I observed a significant increase in theta power at Cue+1 in the Incongruent condition compared with Cue+1 in the Unrelated condition. One interpretation of this result is that this increased theta reflects a prediction error signal triggered by the unexpected category at Cue+1. However, a second significant cluster revealed a significant drop in theta between Cue and Cue+1 in the Unrelated condition. This suggests that the relative increase in theta in the Incongruent condition may be partly driven by decreases in theta power in the Unrelated condition, rather than increases in the Incongruent condition per se. Importantly, a reduction in theta at Cue+1 in the Unrelated condition is consistent with the observed U-shaped function of schema congruency. Previous research has linked theta activity to enhanced memory encoding (Joensen et al., 2023; Kota et al., 2020; Lin et al., 2017), and thus, the decrease in theta at Cue+1 in the Unrelated condition may help to account for poorer memory performance in the Unrelated condition relative to the Congruent and Incongruent conditions. This finding aligns with our behavioural findings, providing evidence that the effects observed at retrieval are indeed driven by neural processes that are present at encoding. However, precisely why theta would decrease between Cue and Cue+1 in the Unrelated condition remains unclear.

One possibility is that the decrease in theta power is caused by sequence position. In this experiment, Cue+1 is always the first or second item in the sequence, and never first or last. Research into serial order retention consistently shows primacy and recency effects whereby memory is enhanced for the initial and final items in a sequence (R. N. A. Henson, 1998; Page & Norris, 1998). These effects are often attributed to the increased allocation of attention to first and last items (Azizian & Polich, 2007; Morrison et al., 2014; Oberauer, 2003). Notably, the upregulation of attentional processes has been associated with increases in theta activity (Fiebelkorn & Kastner, 2019; Keller et al., 2017; McDermott et al., 2017). Therefore, the decrease in theta at Cue+1 may reflect decreases in attention to central items in the sequence. Importantly, research suggests that subjects may pay greater attention to stimuli that are known to predict important outcomes (Le Pelley et al., 2016), which may explain why this reduction was not observed in the Congruent condition. Taken together, these considerations suggest that attentional confounds may complicate direct comparisons between Cue and Cue+1. As such, the Incongruent > Unrelated contrast at Cue+1 may provide a more informative test of prediction error in the present paradigm.

The MVPA analyses in this chapter were aimed at providing evidence of pre-stimulus reactivation of predictable categories. The analyses provided mixed results, but in general they failed to provide conclusive evidence of pre-stimulus category-specific activity that was exclusive to the predictable Congruent condition. The time-resolved decoding analysis confirmed that categories were decodable from the MEG signal in both the Congruent and Unrelated conditions when training on the localiser and testing on the main temporal order memory task. The time-resolved MVPA revealed a significant peak in decoding approximately 250 ms after stimulus onset in both conditions. Notably, however, when peaks were derived at the participant level (i.e., taking each participant's individual peak regardless of when it occurred rather than averaging timepoint-by-timepoint) decoding performance was significantly higher in the Unrelated condition compared to the Congruent condition. Precisely why this pattern emerged isn't entirely clear, but it may align with my previous hypothesis that predictions may extend multiple steps into the future.

Under such a model, expected items are not only predicted at the moment they are about to occur, rather, all upcoming stimuli are partially activated, with more proximal items being represented more strongly than distal items. One possible interpretation of the reduced decoding performance in the predictable Congruent condition is that the predictability of upcoming categories led to their partial activation during the perception of the current category. If multiple category representations are simultaneously active under predictable circumstances, this overlap may reduce classifier performance when the model is trained to discriminate only a single category. Although this analysis alone lacks the specificity to firmly support this conclusion, it nonetheless provides some evidence that predictions are occurring multiple steps into the future.

However, it is worth noting that the pre-stimulus decoding analysis provided no evidence that within-schema categories that are not next could be decoded, raising doubt as to whether representations concerning multiple future categories are indeed active. Notably, however, the pre-stimulus decoding analysis did not account for whether a within-schema category had already occurred or whether it was still upcoming. The presence of categories that had already occurred, and thus would no longer be predicted, may have masked the emergence of decoding for within-schema categories. As such, another potential avenue to explore this hypothesis is to investigate decoding of seen vs. upcoming within-schema items. It may be the case that within-schema categories can be decoded provided the category has not already occurred.

Both the time generalisation analysis and the pre-stimulus decoding analyses were aimed at investigating pre-stimulus category decoding. Permutation tests performed on the time generalisation matrix revealed a significant cluster of decoding in the pre-stimulus period. Interestingly, this significant cluster indicated below baseline classifier performance suggesting that the classifier was actively predicting the negative class. This is a surprising result as neuroimaging research tends to show that reactivated representations are usually analogues of the representations present during perception. Nonetheless, a negative signal still contains information regarding what is coming next. Importantly, however, this signal was present in both the Congruent and Unrelated conditions. This is unexpected as the upcoming category should not be predictable in the Unrelated case, indicating that this pre-stimulus decoding may not be a true predictive signal. In the pre-stimulus decoding analysis, I broke the classifiers down based on whether they were being applied to an epoch when the category it was trained to decode was Next, whether it was Within-schema but not next, or whether it was being applied to an Out-of-schema epoch. Crucially, this revealed that the negative effect observed in the time generalisation analysis was only present when (A) the classifier was applied to an epoch when the category it was trained to decode was Next, and (B) when the classifier was trained on correctly labeled data, random labeling resulted in baseline performance. These findings suggest that the negative effect in the pre-stimulus period is indeed category-specific and is not simply an artifact that is present when the classifier is applied to any pre-stimulus epoch. However, given that the effect is present in both the Congruent and Unrelated conditions, further analyses are required to determine if this is a true predictive signal.

Given that this pre-stimulus effect is present in the Unrelated condition it is worth considering whether the Unrelated condition was as unpredictable as intended. Although participants were told that the Unrelated sequences (referred to as *yellow sequences* to participants) were completely random, they were in fact tightly controlled in order to prevent them from being too similar to the schemas and to ensure that all the foils that were required for the test phase were present. As such, transition probabilities between certain categories in the Unrelated condition were likely more common than others. Given this, there is a possibility that some degree of statistical learning may have taken place. For example, there are certain transitions that never occur in the Unrelated condition. This is because if a transition occurs in the Congruent condition, for example, *face-building,* we don't want this same transition to occur in the Unrelated condition because that could facilitate performance in the Unrelated condition via retrieval-based strategies. If *face-building* occurred in an Unrelated sequence, but *face-building* was also a transition in one of the schemas, then the application of the schema at the point of retrieval could facilitate performance in the Unrelated condition. Therefore, in the Unrelated case, seeing a *face* guarantees that the next category cannot be a *building.*

Likewise, the precise categories that were included in an Unrelated sequence were constrained by the categories that needed to be tested in the other two sequences. For example, imagine a block containing a red, a blue, and a yellow sequence where the red and blue schemas were *face-building-dog-car* and *furniture-tool-insect-plant* respectively. Assuming that the Cue was the second item in both the red and the blue sequence, this would mean that a *dog, car, insect,* and *plant* would need to be tested during retrieval. However, the red and blue sequences cannot provide the appropriate foils for one another, as they do not contain overlapping categories. As a result, the Unrelated sequence is forced to contain a *dog, car, insect,* and *plant,* in order for the necessary foils to be available at test. The order of these categories can vary, but the precise categories themselves are constrained. Moreover, it is noteworthy that in any given sequence, categories never repeat (i.e., there will only ever be one *face* in any single sequence). As such, once you have seen a given category, it can be safely discarded from the list of categories that are likely to come next. Together, these different factors may have combined to allow a degree of statistical learning in the Unrelated condition, facilitating pre-stimulus predictions. Future analyses will be aimed at addressing this hypothesis by partitioning classifier performance in the Unrelated condition based on the transitional probabilities between different categories. In doing so it may be possible to show whether the pre-stimulus activity in the Unrelated condition is dependent on a transition between categories that is statistically highly probable.

However, it is worth noting that peak pre-stimulus discriminant values were highly comparable between the Congruent and Unrelated conditions. This indicates that evidence in favour of the negative class was equivalent across conditions. While some degree of prediction may be possible in the Unrelated case, these predictions should still be weaker than those in the Congruent case, where upcoming categories are fully predictable. The finding that evidence for the negative class is comparable in both conditions therefore raises doubts concerning the interpretation that transitional probabilities could be the factor driving significant decoding in the Unrelated condition.

Previous research (Hetenyi et al., 2024) suggests that pre-stimulus predictions can manifest in the spectral characteristics of a decoding time series. Specifically, they found evidence for increased alpha power in the pre-stimulus period when a category was predictable compared to when classifiers were trained on random labels. To test whether a similar effect emerged in the present paradigm, I computed spectral power from the pre-stimulus decoding time series (-750 to 0 ms). I found that this alpha effect was only observed in the Unrelated condition where upcoming events were uncertain, and not in the predictable Congruent condition. Notably, the cues used in previous research were not 100% predictive of upcoming stimuli. In Hetenyi et al. (2024) predictive cues only predicted upcoming stimuli with 75% accuracy. As such, a possible interpretation of the present finding is that the alpha power signal is not a marker of prediction per se, but a marker of uncertainty. As such, future research may benefit from investigating how this alpha signal relates to the degree of predictability of upcoming events. One way of achieving this with the present data may be to break down the data in the Unrelated condition based on transitional probabilities between categories. If statistical learning is indeed taking place in the Unrelated condition, and this alpha signal is a marker of uncertainty, this may predict that the alpha effect should be stronger for less probable transitions.

Another explanation for why the present study did not see this alpha effect in the predictable Congruent condition may lie in differences in experimental design. Hetenyi et al. (2024) used a design where a specific shape category was predicted by a preceding auditory cue. Although participants were not explicitly told that the auditory cue was predictive, it is likely that participants gained awareness of this. Moreover, the pre-stimulus period was 50% longer in Hetenyi et al. (2024) (1500 ms vs. 1000 ms) than in the present study and participants performed a simple shape discrimination task that may have been less cognitively demanding than the paradigm in the present study. For example, participants were not under the instruction that they needed to remember extended sequences of stimuli for a later memory test. As such, it is plausible that participants in the Hetenyi et al. (2024) study were able to engage in a degree of conscious mental imagery that was less feasible in the present paradigm. Although conscious mental imagery is still a form of prediction, it is worth noting that predictive coding theories propose prediction as a general mechanism of cognitive processing that occurs with or without conscious awareness. As such, further research is required to determine whether these pre-stimulus predictive effects are a general mechanism that occur beyond cases where participants are engaging in conscious imagery.

I also considered whether predictive signals could manifest in the pre-stimulus period but be transformed in such a way as to not be reflective of the representations present during stimulus perception. To test this hypothesis, I ran a cross-validation analysis on the pre-stimulus periods during the main temporal order memory task. Interestingly, I found that representations were decodable during the pre-stimulus period. When the data were labeled based on the upcoming category, above chance decoding was apparent in the Congruent but not Unrelated condition and was especially apparent in the first half of the pre-stimulus period. Notably, however, this was likely due to a confound regarding a category's position in the sequence. In the Congruent condition, the upcoming category always occurs in the same position in the sequence. For example, in the sequence *face-building-dog-car*, *face* always comes first. This is not the case in the Unrelated condition where a category can appear in any position. As such, the above chance decoding in the Congruent condition, that was not present in the Unrelated condition, could be due to classifiers detecting position information rather than category information. This hypothesis was supported by a further analysis that revealed that when the pre-stimulus periods were labeled in terms of sequence position, above chance decoding was also observed in the Unrelated condition. Although this analysis did not provide evidence of pre-stimulus category reactivation, it is interesting that brain activity appeared to track sequence position, with this information being most prominent in the first half of the inter-stimulus interval. Moreover, the negative category-specific signals observed in the time generalisation and pre-stimulus decoding analyses indicated that decoding accuracy for category information was greatest closer to stimulus onset, in the second half of the interval. If future analyses can confirm the nature of the pre-stimulus category-specific signal, it may suggest that during the period between two events in a sequence the brain initially processes position information, which is then taken offline and replaced with information regarding the upcoming category.

Overall, whether pre-stimulus predictions were occurring in the present paradigm remains unresolved. However, it should also be noted that such predictions may simply not be necessary for schema effects to emerge here. Prior knowledge might only be engaged at the point of comparison with sensory input, rather than in advance. In this scenario, sensory inputs would first activate stimulus-specific representations in the brain, which could then be compared against prior knowledge as needed. Although this process would likely slow down cognitive processing and be less adaptive than a system that pre-activates relevant predictable knowledge, it remains a plausible hypothesis. Unfortunately, the present design does not provide the data required to test this potential null hypothesis.

In sum, this experiment successfully replicated the U-shaped function of schema congruency observed in Experiment 4. This indicates that the incongruency effect induced by categorical violations is replicable across experiments. This incongruency effect was also accompanied by evidence of prediction error signals manifesting as increases in theta power approximately 250 ms after stimulus onset on schema-violating Cue+1 trials. I also observed that category, sequence position, and uncertainty information was decodable from the pre-stimulus time series. However, the presence of decoding in both the Congruent and Unrelated conditions indicates that further analyses are required to determine whether the category-specific decoding does indeed reflect a genuine predictive signal.

5 General Discussion

The experiments presented in this thesis were aimed at investigating predictive processing accounts of schema effects in episodic memory. In pursuit of this, three core questions were addressed: (1) Can a U-shaped function of schema congruency be observed in a temporal order memory paradigm, where schemas can facilitate temporal predictions of upcoming events? (2) If so, is it possible to decode these predictions from the pre-stimulus period of an MEG signal, prior to the actual occurrence of a predicted category? (3) When these predictions are violated, are such violations associated with prediction error signals in the brain?

Across four behavioural experiments and one MEG experiment, it was demonstrated that the U-shaped function could indeed be observed. However, while a congruency advantage was present in all experiments, the presence of an incongruency advantage was reliant on the form of violation. Sequential violations, where the violation involved reordering the existing categories in a sequence, did not produce an incongruency advantage. In contrast, categorical violations, where the violation involved the introduction of a category that was completely unexpected for a given context, did elicit a memory advantage for incongruent information. Subsequent MEG analyses revealed that these schema effects were associated with alterations in theta power, specifically in the Unrelated and Incongruent conditions. However, further analyses are needed to establish whether these differences in theta are primarily driven by prediction error or by attentional modulations related to serial order. Finally, although category-specific pre-stimulus decoding was observed, its presence in both the Congruent and Unrelated conditions raises important questions about its interpretation. At present, it remains unclear whether these signals genuinely reflect schema-based predictions, and further analyses are required to resolve this issue.

5.1 Chapter Summaries

5.1.1 Chapter 2 - Schema Effects in Temporal Order Memory

Chapter 2 aimed to investigate whether schema effects could emerge in a temporal order memory paradigm in which the chronology of events formed a core aspect of the schema. Participants were pre-trained on the schematic structure of six-item or four-item sequences in which the schema related to the order of semantic categories. In a subsequent temporal memory task, participants had their memory tested for novel sequences constructed of previously unseen exemplars. Crucially, these sequences could be Congruent, Incongruent, or Unrelated to the pre-learned schemas at the category level. Animacy judgements were measured during sequence encoding to provide a behavioural measure of prediction and prediction error. At test, participants were cued with an item from a sequence and were required to recall the following two items (Cue+1 and Cue+2) in a four-alternative forced-choice task.

Experiment 1 provided two key findings. First, animacy judgements were faster and more accurate in the Congruent condition relative to the Unrelated baseline, indicating that participants could use their prior knowledge of sequence structure to facilitate performance. Moreover, at Cue+1, where the violation occurred in the Incongruent condition, animacy judgements were slower and less accurate in the Incongruent condition than in the Unrelated baseline, providing a behavioural measure of prediction error. Second, at test, a memory advantage was observed for category-level information at Cue+1. Notably, however, there was no effect of incongruency and no effects whatsoever at the exemplar level.

This Experiment provided the first evidence that schema effects, at least congruency effects, could indeed be indexed in a temporal order memory paradigm and when preceded by evidence of predictions during encoding. Interestingly, the lack of an incongruency effect at retrieval indicates that measures of prediction error could be present at the point of encoding, but not be followed by evidence of enhanced learning at test. This suggests that the presence of any prediction error is not sufficient to support learning and that the precise nature of the prediction error bears importance, a fact that many models of learning fail to take into account. I will return to this point in the implications section. Notably, overall performance in Experiment 1 was poor, with approximately 50% of participants failing to meet the inclusion criteria.

Experiment 2 was designed to build upon Experiment 1 in three key ways. First, to improve overall performance, sequence length was reduced from six items to four. This manipulation was effective, and therefore four-item sequences were used in all subsequent Experiments. Second, it aimed to replicate the congruency effect observed in Experiment 1. And third, it attempted to elicit an incongruency effect by introducing controls aimed at increasing the surprise induced by Incongruent trials (e.g., congruent filler blocks, pseudo-randomisation of Incongruent blocks). Experiment 2 successfully replicated the core findings from Experiment 1. A congruency effect was present at the category level at Cue+1 and this was preceded by evidence of predictions and prediction errors at encoding in terms of animacy judgement performance. Notably, an incongruency effect was still not present, indicating that the additional controls and increase in overall performance were not sufficient to induce the effect. This lack on an incongruency effect was further addressed in Experiments 3 and 4.

One limitation with the congruency effects observed in this chapter was that they could potentially be explained in terms of retrieval-based strategies. While the use of schemas at the point of retrieval is indeed an important advantage of possessing schematic knowledge, neurocognitive models of schema effects propose that schemas also modulate encoding. An exploratory analysis of the foils used at retrieval suggested that retrieval-based strategies may not fully account for performance in the Congruent condition. Nonetheless, the demonstration of a category-level incongruency effect or any schema effect at the exemplar level was required to clarify whether encoding-based processes were indeed at play.

Finally, given that prediction errors at encoding did not enhance learning for incongruent items, I aimed to clarify what participants were actually predicting. Although memory was later tested at the category and exemplar levels, responses at encoding were assessed only at the animacy level. This raised the question as to whether participants were predicting specific categories, or merely the animacy of the upcoming category. An exploratory analysis indicated that predictions were primarily occurring at the animacy level, a finding that may help to explain why incongruency effects were not observed at the category and exemplar levels at test.

5.1.2 Chapter 3 - The Influence of Sequential vs. Categorical Violations

The primary goal of Chapter 3 was to investigate the circumstances under which incongruency effects could emerge in a temporal order memory paradigm. This was approached in two ways. First, Experiment 3 replicated the design of Experiment 2, however, each schema was constructed of unique categories that did not overlap between schemas. For example, the schemas could be *face-building-dog-car* and *furniture-tool-plant-insect,* rather than *face-building-dog-car* and *building-car-face-dog.* It was hypothesised that this may reduce interference between the two schemas, as any given category only ever predicted one other, which would facilitate stronger predictions and in turn larger prediction errors. Experiment 4, however, attempted to induce an incongruency advantage by altering the form of surprise altogether. Rather than incongruency being defined as a reordering of existing categories, incongruency now involved the introduction of a category that was completely unexpected for that context. For example, if the Congruent sequence was *face-*

building-dog-car, the Incongruent sequence could be *face-insect-tool-car.* The hypothesis here was that participants may be predicting multiple upcoming categories, including those several steps into the future. If this were the case, incongruency effects may not emerge for the temporal reordering manipulation because, although the categories occur in the wrong order, they are all still partially predicted, weakening any prediction error. By introducing a completely unexpected category, Experiment 4 could potentially generate a prediction error of sufficient magnitude to trigger enhanced learning of Incongruent trials.

The principal finding from this chapter was that incongruency effects were only observed for categorical, and not sequential violations. This provides evidence that during sequence perception, upcoming categories are predicted not one-by-one, but in parallel. Within such a framework, temporal order can still be represented as the strength of each category's activation, with proximal items being represented more strongly than distal ones. However, the partial activation of multiple upcoming items results in weak prediction errors when their order is changed, thereby limiting the emergence of incongruency effects for sequential violations. Crucially, the presence of an incongruency effect in Experiment 4 provided strong evidence that encoding-based mechanisms were contributing to the schema effects observed in this paradigm.

Also of note in this chapter were exploratory analyses showing that behavioural prediction errors at encoding did not influence the presence of an incongruency effect at test. Behavioural prediction errors at encoding appeared to be driven primarily by violations of expected animacy class (animate vs. inanimate) rather than violations of expected category. However, on some trials the violation occurred only at the category level, raising the question as to whether the lack of an animacy violation on some trials was weakening the incongruency effect at test. To test this, differences between the Incongruent and Unrelated conditions at the category level at Cue+1 were broken down by whether the violation at encoding occurred at both the animacy and category level, or at the category level alone. The size of the Incongruency effect did not differ between these cases. This suggests that the factor influencing behavioural prediction errors at encoding, namely, animacy-level information, had no bearing on retrieval performance in the Incongruent condition.

Moreover, two further models respectively used encoding-based accuracy and response time as predictors of Incongruent performance at test, with violation type (sequential vs. categorical) included as a predictor in both models. While violation type predicted Incongruent performance—Incongruent performance was higher for categorical than for sequential violations—neither encoding-based accuracy nor encoding-based response time predicted retrieval-based performance. Collectively, these exploratory analyses further emphasised that the incongruency effect in Experiment 4 was specifically driven by the categorical violation, and not by violations of animacy or by explicit behavioural errors during sequence encoding.

5.1.3 Chapter 4 - The Neural Dynamics of Schema Effects via Magnetoencephalography

Chapter 4 was aimed at investigating the neural dynamics underlying the schema effects observed in Chapters 2 and 3. To do this, it combined the behavioural paradigm from Experiment 4 with magnetoencephalography (MEG) and multivariate decoding methods. The first objective was a behavioural replication of the U-shaped function of schema congruency observed in Experiment 4, whereby categorical violations elicited an incongruency advantage at the category level. In addition, I investigated whether Incongruent trials were accompanied by neural signatures of prediction error, such as alterations in theta power. Multivariate decoding was then used to try and provide evidence of pre-stimulus predictive processing during the presentation of sequences with a predictable structure. Specifically, it was hypothesised that pre-stimulus decoding would be possible in the Congreunt condition, but not the Unrelated condition. Together, these findings would provide strong evidence that the schema effects observed throughout the present thesis were indeed driven by predictive coding mechanisms.

The MEG experiment successfully replicated the schema effects observed in Experiment 4. During retrieval, accuracy was higher in both the Congruent and Incongruent conditions relative to the Unrelated baseline at the category level at Cue+1. This finding provides good evidence that categorical violations reliably elicit an incongruency advantage, thereby facilitating the emergence of a U-shaped function of schema congruency in a temporal order memory task. Notably, the MEG experiment did not include animacy judgements during sequence encoding, indicating that these effects can emerge in the absence of an encoding task.

When running time-frequency analyses on Cue, Cue+1 and Cue+2 trials at encoding there were modulations of theta activity. When contrasting the Incongruent and Unrelated conditions at Cue+1, there was significantly higher theta in the Incongruent condition from approximately 250 ms after stimulus onset. It is at Cue+1 in the Incongruent condition that the violation occurs, suggesting that this theta increase may be a marker of prediction error in the brain. Notably, however, when comparing Cue and Cue+1 in the Unrelated condition there was a significant decrease in theta power at Cue+1. This suggests that a decrease in theta at Cue+1 in the Unrelated condition may have, at least in part, contributed to the signal observed in the Incongruent > Unrelated comparison, raising questions as to whether it was a true prediction error signal. Nonetheless, this decrease in theta power in the Unrelated condition is in line with previous findings that suggest greater theta power is associated with enhanced encoding (Joensen et al., 2023; Kota et al., 2020; Lin et al., 2017). The lower theta power observed in the Unrelated condition at Cue+1 may explain why memory retention was worse in the Unrelated condition as compared to the Congruent and Incongruent conditions.

MVPA analyses were conducted to determine if category-specific representations were reactivated during the pre-stimulus period in predictable schema-congruent sequences. Initial time-resolved decoding, where classifiers were trained and tested on each respective time point, revealed that the categories could be decoded after stimulus onset, with a peak decoding time of approximately 250 ms. Interestingly, peak decoding values were higher in the Unrelated than in the Congruent condition. Time generalisation analysis between -400 to 1000 ms provided the first evidence of pre-stimulus decoding. A significant cluster was observed in the pre-stimulus period of the test set when the classifiers were trained on times when the stimuli were on screen. Notably, however, decoding values were significantly below chance, suggesting the classifier was actively predicting the negative class in the pre-stimulus period. Moreover, pre-stimulus decoding was present in both the Congruent and Unrelated conditions, despite categories being unpredictable in the Unrelated condition. This pre-stimulus decoding in the Unrelated condition raised questions as to whether the signal was genuine category-specific decoding or a more general artifact of applying classifiers to pre-stimulus activity. To address this issue, I assessed raw discriminant values of classifiers applied to epochs where a given category was Next, Within-schema, or Out-of-schema. Interestingly, this analysis indicated that the pre-stimulus decoding was only present when a classifier's target category was Next, indicating that the pre-stimulus decoding was indeed category-specific. However, why this activity was present in both the Congruent and Unrelated conditions remains unresolved.

Next, I investigated whether pre-stimulus predictions were represented in the frequency characteristics of the pre-stimulus decoding time series. Previous research suggests that pre-stimulus predictions may manifest as increases in alpha power in the decoding time series, relative to classifiers trained on random labels (Hetenyi et al., 2024). Interestingly, I observed a similar increase in alpha power in the pre-stimulus decoding time series, but only in the Unrelated condition and not in the predictable Congruent condition. One interpretation of this result is that elevated alpha power does not reflect category-specific predictions, but is instead a marker of uncertainty.

Finally, cross-validation analyses of the pre-stimulus periods in the main temporal order memory task revealed significant above chance decoding throughout the time series. Notably, whether this decoding was significant in the Unrelated condition depended on how the data were labelled. When labelled in terms of the upcoming category, decoding was significant in the Congruent case but not in the Unrelated case. However, when labelled in terms of sequence position, decoding was observed in both the Congruent and Unrelated conditions. Given that upcoming category and sequence position correlate in the Congruent condition, this indicates that the decoding was primarily driven by sequence position rather than the upcoming category. Interestingly, this decoding was most prominent during the first half of the pre-stimulus period. This is in contrast to the category-specific decoding observed in the previous analyses that occurred during the second half of the pre-stimulus period. This temporal dissociation may indicate that during the inter-stimulus interval, sequence position information is initially activated, before being replaced with category information closer to stimulus onset.

Overall, these analyses demonstrated that information is decodable during the pre-stimulus period. Specifically, the results suggest that category information, sequential information, and markers of uncertainty may all be present prior to stimulus onset. However, further analyses are required to determine whether the category-specific decoding is indeed a genuine predictive signal.

5.2 Implications and Future Directions

This thesis used a temporal order memory paradigm in which participants were trained on the schematic structure of a sequence. This paradigm was designed to enable participants to use their prior knowledge of sequence structure to make temporal predictions of upcoming events. Demonstrating schema effects in this context has important theoretical implications, as schema effects are generally thought to arise due to predictive processing in the brain. Predictive processing theories of schema effects suggest that schema effects emerge due to a comparison of sensory experience with expected inputs, a process that in turn functionally modulates memory encoding. A central question raised by this framework is precisely how and when this comparison process occurs. Specifically, does predictive processing entail the pre-activation of expected inputs prior to their actual occurrence? Or does perception unfold first, with schematic knowledge being brought online further down the perceptual hierarchy? If schemas are routinely employed to predict future events prior to their occurrence, schema effects should be observed in a temporal order memory paradigm when the chronology of events forms a core aspect of the schema. Moreover, the processing of information related to the schematic structure of the sequences should be observed both when stimuli are actively online but also during pre-stimulus periods.

The presence of schema effects in the present thesis is in line with a model by which schemas facilitate temporal predictions of the future. In the behavioural studies, I observed that schemas facilitated animacy judgements both in terms of accuracy and response time indicating that prior knowledge was indeed being used to predict upcoming items. Furthermore, I observed subsequent memory advantages at retrieval. These findings are in line with previous research that suggests that prior knowledge can be used to facilitate perceptual processing (M. Smith et al., 2021; M. E. Smith & Loschky, 2019) and boost subsequent memory (Gilboa & Marlatte, 2017; Greve et al., 2019; Huang et al., 2023; Ortiz-Tudela et al., 2024; Raykov et al., 2019; van Kesteren et al., 2012). Together, these results show that expected events were identified more accurately, responded to more rapidly, and recalled more reliably, emphasising the role that schemas play in promoting efficient perceptual and mnemonic processing, likely via pre-stimulus expectations.

5.2.1 Sequential vs. categorical violations

The above findings are consistent with the view that schemas are used to preempt expected sensory inputs, and that this can modulate subsequent memory retention. However, the findings observed in the schema-incongruent condition have important implications for how predictions of future events may be occurring. Across four behavioural experiments and an MEG experiment, I observed that incongruency effects were observed for categorical violations, but not for violations of sequential order. Incongruency advantages are believed to be a consequence of prediction errors whereby actual events conflict with expectations. Both sequential and categorical violations conflict with what should be predicted to occur, raising the question as to why memory advantages are only seen for categorical violations.

One model of how predictions occur is that the brain maintains a model of the entire sequence which is rolled out sequentially (Daw & Dayan, 2014). Here, each category would be individually pre-activated during the pre-stimulus period before it occurs. Alternatively, predictions may occur multiple steps into the future. Rather than single categories being predicted as they occur, multiple upcoming categories may be represented at any given time (Dayan, 1993). But how might this account for the categorical vs. sequential distinction? If multiple upcoming items are predicted in the case of sequential violations, then although the categories occur in the wrong order, the incongruent category is still somewhat predicted, weakening any prediction error. However, for categorical violations, the incongruent category is not predicted to occur at any point in the sequence, allowing for a strong prediction error that can drive new learning.

Notably, predictions occurring multiple steps into the future raises the question as to whether predictions are indeed temporal in nature. If all categories are equally pre-activated, how is temporal order represented? If temporal order is not represented, equal pre-activation of all expected categories makes sequences with sequential violations indistinguishable from schema-congruent sequences. For example, if a Congruent sequence was *face-building-dog-car,* and an Incongruent sequence was *face-car-dog-building,* then with all categories equally predicted, the Incongruent sequence would still be fully in line with expectations. If this is indeed the case, this model should predict congruency advantages despite temporal reordering, yet no advantage was observed.

One possible explanation for the above problem is that congruency advantages in the Congruent condition are purely driven by retrieval-based processes. If encoding-based processes do not contribute to the congruency effects observed here, then the lack of a “congruency” advantage for sequences with sequential violations may simply reflect the ineffectiveness of retrieval-based strategies in that context. However, my exploratory analyses in Chapter 2 suggested that retrieval-based strategies may not fully account for the observed congruency advantages, and so it is worth considering other explanations.

An alternative explanation for the absence of an incongruency advantage for sequential violations is that temporal order information is, in fact, represented. If multiple upcoming items are predicted at any given time, temporal order information could be represented as the relative strength of each category’s activation. For example, more proximal items may be activated more strongly than more distal ones. If this were the case, this could account for why congruency advantages are observed in the Congruent but not Incongruent conditions for sequential violations, but also why new learning could be perturbed by weakened prediction errors. In the case of sequential violations, items that are upcoming but not next will still be partially activated, this partial activation weakens the prediction error triggered by the violation, and therefore no incongruency benefit is observed. Taken together, these findings suggest that the data in the present thesis may be most in line with a model where predictions are occurring multiple steps into the future, with graduated levels of pre-activation dependent on the proximity of an upcoming category.

This model, in which incongruency effects are absent for sequential violations due to the prediction of multiple upcoming items, should be tested in several keys ways. First, although multiple future events may be predicted, this process cannot extend indefinitely. In a model where more distal events are represented less strongly, an event will eventually be far enough in to the future such that its occurrence is not predicted strongly enough to suppress the prediction error required to drive new learning. In other words, sequential violations should elicit incongruency effects provided the violation involves a category that is sufficiently far into the future. One problem with testing this hypothesis with the current paradigm is that, as observed in Experiment 1, extending the sequences even from four items to six results in a severe reduction in overall performance. Nonetheless, if a paradigm that facilitated greater performance with longer sequences could be developed, this could be an important test of the above hypothesis as to why incongruency effects were not observed for sequential violations. One approach could be to use a longer, recursive sequence that forms a loop (e.g., Tarder-Stoll et al., 2024). For instance, the loop could consist of ten categories and serve as the schema that participants are required to pre-learn. During the main temporal order memory task, participants would only need to encode and retrieve four-item chunks from this larger loop. By introducing sequential violations using categories that are either proximal or distal within the loop, it may be possible to reveal whether sequential violations elicit incongruency benefits depending on the proximity of the violating category.

A second way in which this model could be tested could be to introduce a narrative structure to the sequences. The sequences used in the present paradigm were highly abstract, with no inherent narrative structure. The abstract nature of the sequences was primarily employed due to the aim of combining the paradigm with multivariate decoding in the MEG experiments. Narratives are defined by their causal structure across time, and narratives likely play an important role in linking causes to their outcomes in complex environments (Chen & Bornstein, 2024). Moreover, narratives have frequently been shown to enhance overall memory retention (Chen & Bornstein, 2024; Cohn-Sheehy, Delarazan, Reagh, et al., 2021; Cohn-Sheehy, Delarazan, Crivelli-Decker, et al., 2021; Lee & Chen, 2022). Importantly, the casual structure of real-world events may place greater constraints on the temporal order in which events should logically occur. Hence, the lack of a narrative in the present sequences may have resulted in weak priors regarding the temporal ordering of events, questioning the ecologically validity of the finding that incongruency effects are not observed for sequential violations. As such, future research should introduce a narrative structure to the sequences to investigate whether the sequential vs. categorical distinction still holds.

5.2.2 Not all violations drive learning

In this thesis, expectations were violated at several different levels of abstraction: at the category level, animacy level, or both. Violations could involve the re-ordering of existing categories, or the introduction of a category that was completely unexpected for a given context. Moreover, violations could also be accompanied by behavioural errors, such as incorrect animacy judgements or increased response times. Notably, despite these various forms of schema-incongruence, enhanced learning was only observed under very specific circumstances. Specifically, incongruency advantages were only observed at the category level for categorical violations, with explicit behavioural errors bearing no influence on this effect.

Classic models of learning typically only account for the magnitude of prediction error (Rescorla & Wagner, 1972). Later models incorporated modelling of attention, which could account for factors such as stimulus familiarity (Pearce & Hall, 1980), however, they often still failed to capture nuances such as contextual and qualitative differences in the nature of prediction errors. More recent Bayesian models have made advances in this area, modelling prediction error as the divergence between probability distributions, allowing for the manipulation of not just magnitude but also the precision of priors and sensory evidence (R. N. Henson & Gagnepain, 2010). By modelling prediction across multiple levels of representation (e.g., perceptual, semantic, episodic/contextual) these models may come closer to accounting for the ways in which predictions can qualitatively vary. Nonetheless, the results presented in this thesis further highlight the need to determine precisely what features of a prediction error are required to drive new learning.

Recent work studying individuals from multiple age groups has likewise brought attention to this issue. Nolden et al. (2025) pre-trained children, younger adults, and older adults with sequences of paired-associates where pairs of items predicted each other but pairs could be presented randomly throughout the sequence. Following pre-training, participants would view sequences in which new items were either placed between (non-violating) or within pairs (violating). In subsequent recognition and association tests, there was no memory benefit for violating items in any of the age groups. Based on these findings, Nolden et al. (2025) concluded that prediction error may not always benefit episodic memory as consistently as some theories suggest. Comparing the paradigm used by Nolden et al. (2025) with those used in the present thesis may help to elucidate what factors may be critical for the emergence of memory advantages from violating events.

Both Nolden et al. (2025) and the present paradigm presented participants with sequences in which items predicted upcoming events. However, the present paradigm imposed stronger constraints on which items could appear within a more limited contextual frame. Whereas Nolden et al. (2025) presented participants with blocks of extended sequences of 144 items in which pairs of items predicted one another, the present paradigm involved pre-training knowledge that could then be applied to novel sequences that were much more constrained in length and content. It is possible that when the set of possible categories/exemplars within a context is too broad, this may limit the extent to which violating events can potentiate learning. This suggests that the context in which the predictions and violations are occurring, including the number of expected items and degree to which they are expected to occur, can influence whether a prediction error leads to new learning.

A further distinction lies in the type of learning involved and the way in which it was applied. Nolden et al. (2025) relied on statistical learning, whereas the present research required participants to acquire explicit knowledge of the sequence's schematic structure. Although participants were not explicitly instructed to apply this knowledge, they were nevertheless highly aware of its utility in the memory tests. This raises the possibility that enhanced learning from violations may depend on the relevance of the unexpected event to current behavioural goals. After all, if an event is unexpected but irrelevant to ongoing goals, why expend resources encoding it? If there is no clear goal in a task, violations of expectation may not be perceived as behaviourally relevant, reducing their salience and in turn new learning.

In sum, the above comparison suggests that the extent to which an ongoing context constrains relevant predictions, and the degree to which this context promotes the behavioural relevance of the information that can be violated, must be considered when attempting to experimentally induce incongruency effects.

Finally, of note in the present thesis was a boost in memory retention specifically for information at the level at which the violation occurred. That is, I observed a boost in memory for information at the category level specifically when violations involved a completely unexpected category. Notably, I observed little evidence for a memory advantage at the exemplar level despite this information being central to the goals of the temporal order memory task. This raises the question as to whether prediction error benefits memory not just based on the raw magnitude of the error, but also based on qualitative aspects such as whether the error occurred at the same (or similar) level of representation as the to-be-remembered information. Greve et al. (2019) found incongruency effects for displays shown at test when the violation related to a value rule, suggesting that the memory benefit of an error can generalise to other forms of information, however, Greve et al. (2019) used a recognition test rather than cued recall. Previous research that has used forward cued recall has failed to find a memory benefit for incongruent events following the cue (Frank et al., 2018). Yacoby et al. (2025) found a memory benefit for incongruent items in a 2AFC cued recall task, however, this task did not allow for discernment of performance at different levels of analysis (e.g., category vs. exemplar).

The idea that prediction errors may be sensitive to qualitative factors, such as the representational level of the violating information, implies that violations at different levels of representation should preferentially benefit memory for information at the corresponding level. Future research could test this theory by introducing sequence violations that occur at the exemplar, rather than category, level. If a stronger memory benefit could be observed for exemplar level information in this case it would provide good evidence that the qualitative nature of a prediction error, and not just its magnitude, is important for learning. This finding would have important implications for existing neurocognitive models of schema effects (e.g., SLIMM, van Kesteren et al., 2012) that predict that incongruency should enhance memory for all information, including incidental details, rather than selective enhancement for specific levels of information. Moreover, it would provide an elegant explanation for why exemplar-level schema effects were mostly absent in the present thesis.

5.2.3 Neural markers of memory and prediction error

A core goal of the work in this thesis was to provide evidence of neural markers of predictive coding and prediction error in a context where temporal predictions of future events were possible. Past research has frequently revealed that prediction errors manifest as alterations in theta-band activity (Garrido et al., 2015; Recasens et al., 2018). In Chapter 4, time-frequency analyses revealed modulation of theta across the Congruency conditions. Most in line with a prediction error signal was a significant cluster of activation when contrasting the Incongruent and Unrelated conditions at Cue+1 during encoding. Cue+1 is the point of violation in the Incongruent condition indicating that this increased theta activity may represent a prediction error.

Notably, this increase in theta occurred at approximately 250 ms after stimulus onset. When a violating stimulus is presented, it must first reach the visual cortex before propagating throughout the brain so that different regions can compare it with existing knowledge. As such, it is worth considering how the timing of this theta signal relates to known network dynamics in the brain to establish whether pre-activation is required to account for prediction errors occurring at this time. The increase in theta here seemed to emerge at approximately the same time at which multivariate decoding of category information peaked. However, category decoding was statistically significant as early as 80 ms after stimulus onset. Assuming the onset of significant decoding represents the time at which category-specific information was present in visual cortex, would this leave enough time for a comparison process before a prediction error signal emerged at 250 ms?

With regards to memory recall, research suggests that cued information can reach MTL within ~500 ms and that old-new familiarity signals arise within this same time frame (Staresina & Wimber, 2019). Hippocampal pattern completion is believed to occur between ~500 - 1500 ms after cue onset, triggering conscious recollection. However, research suggests that mismatch signals can be detected in the brain much faster than this. The early mismatch negativity is an ERP response occurring approximately 100 – 200 ms post stimulus. This negativity is a response to unexpected changes in sensory input and is thought to reflect early, pre-attentive change detection in the brain (Heslenfeld, 2003; Kimura et al., 2009; Stefanics et al., 2014). Notably, studies into the visual mismatch negativity (vMMN) frequently use paradigms in which violations are introduced as oddballs within an ongoing pattern. As such, the vMMN may reflect comparisons with information stored in sensory or working memory as an event unfolds, rather than comparisons with long-term memory representations. This contrasts with the present paradigm in which schemas stored in long-term memory would need to be activated based on the given context, a situation in which the mismatch cannot be inferred from recent sensory inputs alone and prior knowledge of the schemas is required. Therefore, the vMMN's early occurrence after stimulus onset may arise for reasons that do not apply for the present temporal order memory paradigm.

The N170 potential has likewise been proposed to be a prediction error signal suggesting rapid error handling in the brain (Baker et al., 2021, 2023; Johnston et al., 2017; Robinson et al., 2018). However, as above in relation to the mismatch negativity experiments, these investigations used paradigms where mismatch signals were based on violations of ongoing motion trajectories or facial expressions rather than comparisons with long-term knowledge (Johnston et al., 2017). Notably, however, recent meta-analyses suggest that the N170 is increased in size for personally familiar faces (Caharel & Rossion, 2021). Although not an error signal, this may indicate that long-term knowledge can be compared with a sensory input to modulate cognition within 200 ms. However, Dobs et al. (2019), found that the perception of familiar faces is likely a consequence of the tuning of early feed forward processes, as there is unlikely to be time for feedback from higher-order brain regions (Dobs et al., 2019; Lamme & Roelfsema, 2000)

Crucially, the two regions that are most important to consider when discussing prediction errors caused by schema violations are the mPFC and MTL. The mPFC is believed to be important for schema monitoring, or the integration of prior knowledge with ongoing encoding, a feature which is shared by multiple models of PFC-MTL interaction (Preston & Eichenbaum, 2013; Robin & Moscovitch, 2017; Takehara-Nishiuchi, 2020; van Kesteren et al., 2012). If the mPFC generates pre-stimulus expectations based on prior knowledge, it could theoretically send these signals to sensory regions allowing for rapid comparisons of sensory input with prior knowledge. However, if expectation signals are not pre-activated, sensory information would need to reach the mPFC before it could be assessed. Additionally, under SLIMM, the mPFC detects whether sensory inputs are congruent with pre-existing knowledge and inhibits the hippocampus under schema-congruent conditions. This implies that sensory information must have reached the mPFC in order for the appropriate inhibition, or lack thereof, to occur. This raises the question as to how quickly sensory information could reach the mPFC, and then potentially the HPC, and whether this could occur within 250 ms of stimulus onset in order to align with the onset the observed theta increase. Notably, research suggests that visual information can reach the prefrontal cortex very rapidly, in under ~100-150 ms (Bellet et al., 2022; Foxe &

Simpson, 2002; Kwon et al., 2021), suggesting that pre-activation may not necessarily be required to account for the onset of this theta signal.

As it is, it remains unclear whether or not the timing of this theta signal could be accounted for by a model in which prior expectations are not pre-activated. Future analyses will attempt to address this issue by performing source reconstruction. By performing analyses within specific regions-of-interest such as the mPFC and MTL, it may be possible to determine at what times different brain regions are able to respond to and represent different kinds of information across the Congruency conditions, and whether prediction errors triggered by comparisons against long-term memory representations can occur in under 250 ms.

It is also worth considering whether the increase in theta in the Incongruent > Unrelated contrast is indeed representative of a genuine prediction error signal. Prediction errors should not be present in the Congruent condition either. However, although numerically higher in the Incongruent condition, the Incongruent > Congruent contrast at Cue+1 did not show a significant cluster for increased theta activity. Moreover, when comparing Cue+1 > Cue in each of the three Congruency conditions, there was a significant decrease in theta in the Unrelated condition and no significant difference in theta in the Incongruent condition. This indicates that increases in the Incongruent > Unrelated comparison may be more so driven by difference in the Unrelated condition rather than increases in the Incongruent condition per se. It remains unclear as to why this decrease would occur at Cue+1 in the Unrelated condition, however, one possibility is that it is related to sequential effects. The present paradigm used four-item sequences, participants were cued with either the first or second item in the sequence and required to recall the following two items. This meant that Cue+1 was always either the second or the third item in the sequence, it could never be the first or last item. Therefore, differences between Cue and Cue+1 may reflect processes related to sequence position with central items being processed differently than beginning or end items.

Research suggests that participants allocate more attention to the first and last items in a sequence, and increased allocations of attention have been associated with increased theta power (Fiebelkorn & Kastner, 2019; Keller et al., 2017; McDermott et al., 2017). Thus, the decrease in theta for central items in the Unrelated condition may be a consequence of reduced attention to items in the centre of a sequence. This may explain why the theta decrease was observed at Cue+1 but not at Cue+2. Cue+2 can either be the third (penultimate) or fourth (last) item in the sequence, meaning that the heightened attention to final items may have resulted in increased theta activity for the subset of Cue+2 trials that occurred at the end of a sequence. In the Incongruent condition, this attentional theta effect may be masked by prediction errors triggered by violating Cue+1 trials. With regards to the Congruent condition, research suggests that participants pay more attention to items that are predictive of upcoming events (Le Pelley et al., 2016) which may explain the absence of a theta decrease. That said, other lines of research suggest that predictive cues, such as those in the Congruent condition, may in fact be associated with decreased theta power (Carmo-Blanco & Allen, 2019; Crivelli-Decker et al., 2018). As such, future source space and topographical analyses will be important for enhancing statistical power, and identifying the brain regions generating these signals, to provide insight into the contributions of predictive and attentional processes to theta modulation.

Another interesting finding in these time-frequency analyses was a lack of any differences between the Congruent and Unrelated conditions. As has been noted, it is unclear from my behavioural results whether the observed congruency effects are partly driven by encoding-based mechanism, or whether they are wholly a consequence of retrieval-based strategies. The present results provide no evidence of a frequency domain marker of enhanced encoding in the Congruent case, questioning whether encoding-based mechanisms are at play here. However, it is noteworthy that this single sensor space analysis may not be specific enough, or may lack the statistical power required, to unveil markers of increased encoding in the Congruent condition. Existing neurocognitive models of schema effects suggest that schema monitoring is localised to specific regions, most notably the mPFC (Preston & Eichenbaum, 2013; Robin & Moscovitch, 2017; Takehara-Nishiuchi, 2020; van Kesteren et al., 2012). Moreover, SLIMM predicts that the influence of schema congruence on neural activity may manifest as differences in mPFC-MTL connectivity. Under schema-congruent conditions, SLIMM predicts that the mPFC should inhibit hippocampal encoding. Whether inhibition should manifest as more or less functional connectivity is unclear, but some variation in connectivity should still be expected between Congruency conditions. As such, future source space analyses are required to fully establish whether or not encoding-based markers of schema-congruent encoding are present.

5.2.4 Pre-stimulus decoding

The present research demonstrates that three different types of information can be decoded from a pre-stimulus period: category information, sequential information, and markers of uncertainty.

Most notably, pre-stimulus decoding of category information was present in both the Congruent and Unrelated conditions raising questions as to whether this is a genuine predictive signal. By breaking the data down according to whether a category was expected Next, was Within-schema, or was Out-of-schema, it was shown that category decoding was only present when a category was Next. This indicates that the decoding of category information is indeed category-specific and not an artifact observed when applying a classifier to any pre-stimulus time series. This finding of decoding in an Unrelated condition highlights the importance of including unpredictable control conditions to ensure pre-stimulus decoding is genuinely driven by predictions and not some other factor. Previous studies that have demonstrated pre-stimulus decoding have shown that distinguishing perceptual categories is possible when they are preceded by different predictive cues, however, there was no clear unpredictable, or fully predictable, control to demonstrate that this decoding only occurred when things were genuinely predictable (Kok et al., 2017).

However, a potential explanation for the pre-stimulus decoding in the present Unrelated condition is that statistical learning is taking place. Although participants were told that the Unrelated sequences were entirely random, they were in fact tightly constrained to prevent the occurrence of schema-congruent transitions in the Unrelated condition and to ensure the foils needed at test were present. Although plausible, a key weakness of this explanation is that peak decoding prior to stimulus onset seemed to be equivalent in both the Congruent and Unrelated conditions. If statistical learning was promoting the decoding of category information in the Unrelated condition, one would expect this signal to at least be weaker than that observed in the Congruent condition, as the Congruent condition is fully predictable whereas the Unrelated condition is at best probabilistic. Notably, pre-stimulus decoding in the Congruent condition was significant at an earlier time point than in the Unrelated condition, potentially suggesting that category information was more robustly available throughout the pre-stimulus time series in the Congruent condition. Future analyses regarding the onset of decoding in the pre-stimulus time series between Congruency conditions may therefore help to untangle whether category information is more prominent in the Congruent condition.

Another key direction for pre-stimulus decoding analyses will be to decode in the frequency domain and to use techniques such as Representational Similarity Analysis (RSA) instead of MVPA. When combining these two approaches previous research has detected decoding in the pre-stimulus time period (Sherman et al., 2022). The present thesis did not include these approaches as it aimed to be more in line with research that has decoded pre-stimulus activity using MVPA approaches (e.g., Hetenyi et al., 2024; Kok et al., 2017). A key advantage of using RSA instead of MVPA is that it may help to distinguish the degree to which representational content is available across different conditions. With MVPA, it is difficult to know precisely what information the classifier is using to make its predictions and why information has led to the hyperplane being placed where it is. Using a correlational approach such as RSA there may be clearer insight into the underlying representational geometry. For example, RSA may help to determine whether decoding is driven by true categorical distinctions or lower-level perceptual features. In sum, RSA could be a powerful complement to MVPA, providing not just evidence that information is present, but also information about the structure of the representations that support successful decoding.

Cross-validation analyses on the pre-stimulus time series during the main temporal order memory task also revealed that sequence position information could be decoded, and provided little evidence that category-specific information was contributing to this effect. The lack of evidence of pre-stimulus decoding of category information using cross-validation indicates that if category information is present in the pre-stimulus period, it is indeed reflective of the pattern of activity observed during stimulus perception rather than being distinct from the perceptual representation is some way. Indeed, this finding would be in line with other studies that have revealed pre-stimulus decoding and have found it to be reflective of patterns of activity observed during perception (Kok et al., 2017; Sherman et al., 2022).

Finally, I was able to observe increased alpha power in a pre-stimulus decoding time series when averaging power across the canonical alpha frequency bands (8 - 12 Hz), a similar effect to that seen in previous studies (Hetenyi et al., 2024). Notably, however, the difference in alpha power between classifiers trained on accurate vs. random labels was not observed in the predictable Congruent condition, but in the unpredictable Unrelated condition. This suggests that the increase in alpha power is a robust phenomena, however, it may not strictly be a predictive signal as previously proposed. Previous research that has observed this effect used predictive cues that were only predictive of an upcoming event in a probabilistic manner, they predicted the upcoming event with only partial accuracy. Given this, and the results in the present work, this may indicate that the effect is instead a marker of uncertainty regarding an upcoming event, rather than a marker of prediction per se. Future research should aim to provide further evidence of this interpretation. One interesting question that arises from this effect being absent from the Congruent condition, is whether the size of the pre-stimulus alpha effect scales with the degree of predictability. Crucially, this proposal could help to account for whether statistical learning is occurring in the Unrelated condition. If this proposed uncertainty signal could be shown to be most prominent for Unrelated transitions that are less probable than others, this could provide some further evidence that statistical learning was indeed taking place and that the alpha signal is indeed a marker of uncertainty.

Together, the findings that category, sequential, and uncertainty signals can be decoded during a pre-stimulus period provides an interesting picture of how cognitive processing may unfold between stimuli during the perception of a predictable sequence. Throughout the inter-stimulus interval (ISI), sequential information is tracked, providing knowledge of position in the sequence. This continues throughout the ISI but is most prominent in the first 500 ms. Following this, category information comes online as stimulus onset approaches. Notably, the degree to which a category is predictable may influence the time course at which it can be decoded, with more predictable categories being decodable at earlier time points. Finally, uncertainty regarding an upcoming category may be coded in the frequency characteristics on the decoding time series, specifically in terms of alpha power. Precisely when this uncertainty signal is most prominent remains an open question that could be addressed with future time-frequency decompositions. Although further analyses are required to fully establish this proposal, the analyses presented in this thesis provide novel evidence that multiple sources of information are present in the pre-stimulus period, offering new insight into the mechanisms that prepare perceptual and mnemonic systems for upcoming input.

5.3 Conclusion

This thesis aimed at investigating the presence of schema effects in a temporal order memory paradigm when the chronology of events formed a core aspect of a schema's content. Observing schema effects under such conditions would provide evidence in favour of a predictive processing account of schema effects in memory, suggesting that prior knowledge is used to make temporal predictions of upcoming events. Across four behavioural experiments and one MEG experiment, I observed that schema-congruent sequences reliably resulted in enhanced memory retention. However, the emergence of a memory benefit for schematically incongruent sequences was dependent on the form of violation. Incongruency effects were observed for categorical, but not sequential violations. I propose that this finding is most in line with a predictive account, whereby prior knowledge is used to predict the category of multiple upcoming items. Under this framework, temporal order information is represented in terms of the strength of activation of an upcoming category, with proximal items pre-activated more strongly than distal ones. The MEG study provided evidence that theta oscillations were modulating prediction error signals and attentional processes. However, further source space analyses are required to refine these findings. During a pre-stimulus period, category, sequential, and uncertainty signals were present in the decoding time series, providing evidence that prior knowledge is used to predict how a sequence is likely to unfold. However, the presence of category-specific decoding in both the Congruent (predictable) and Unrelated (unpredictable) conditions means that further analyses are required to fully resolve whether these are genuine predictive signals.

Overall, this thesis demonstrates that schemas can guide temporal predictions in memory and links these predictions to subsequent memory benefits for schema-congruent and schema-incongruent events. Notably, the MEG findings indicate that theta oscillations mediate predictive and attentional processes, and that pre-stimulus activity carries information about upcoming category, sequence structure, and prediction uncertainty, consistent with the idea that prior knowledge is used to anticipate future events.

References

Azizian, A., & Polich, J. (2007). Evidence for Attentional Gradient in the Serial Position Memory Curve from Event-related Potentials. *Journal of Cognitive Neuroscience*, *19*(12), 2071–2081. https://doi.org/10.1162/jocn.2007.19.12.2071

Baker, K. S., Johnston, P., Yamamoto, N., & Pegna, A. J. (2023). Event-Related Potentials Index Prediction Error Signalling During Perceptual Processing of Emotional Facial Expressions. *Brain Topography*, *36*(3), 419–432. https://doi.org/10.1007/s10548-023-00951-2

Baker, K. S., Pegna, A. J., Yamamoto, N., & Johnston, P. (2021). Attention and prediction modulations in expected and unexpected visuospatial trajectories. *PLOS ONE*, *16*(10), e0242753. https://doi.org/10.1371/journal.pone.0242753

Barron, H. C., Auksztulewicz, R., & Friston, K. (2020). Prediction and memory: A predictive coding account. *Progress in Neurobiology*, *192*, 101821. https://doi.org/10.1016/j.pneurobio.2020.101821

Bartlett, F. (1932). *Remembering: A study in experimental and social psychology.* Cambridge University Press.

Bein, O., Duncan, K., & Davachi, L. (2020). Mnemonic prediction errors bias hippocampal states. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-17287-1

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

Bein, O., Reggev, N., & Maril, A. (2014). Prior knowledge influences on hippocampus and medial prefrontal cortex interactions in subsequent memory. *Neuropsychologia*, *64*, 320–330. https://doi.org/10.1016/j.neuropsychologia.2014.09.046

Bellet, J., Gay, M., Dwarakanath, A., Jarraya, B., Kerkoerle, T. van, Dehaene, S., & Panagiotaropoulos, T. I. (2022). Decoding rapidly presented visual stimuli from prefrontal ensembles without report nor post-perceptual processing. *Neuroscience of Consciousness*, *2022*(1). https://doi.org/10.1093/nc/niac005

Blanchet, S., Gagnon, G., & Bastien, C. (2007). Event-related potential study of dynamic neural mechanisms of semantic organizational strategies in verbal learning. *Brain Research*, *1170*, 59–70. https://doi.org/10.1016/j.brainres.2007.07.024

Bowles, B., Crupi, C., Mirsattari, S. M., Pigott, S. E., Parrent, A. G., Pruessner, J. C., Yonelinas, A. P., & Köhler, S. (2007). Impaired familiarity with preserved recollection after anterior temporal-lobe resection that spares the hippocampus. *Proceedings of the National Academy of Sciences*, *104*(41), 16382–16387. https://doi.org/10.1073/pnas.0705273104

Bramão, I., & Johansson, M. (2018). Neural Pattern Classification Tracks Transfer-Appropriate Processing in Episodic Memory. *Eneuro*, *5*(4), ENEURO.0251–18.2018. https://doi.org/10.1523/eneuro.0251-18.2018

Brehm, L., & Alday, P. M. (2022). Contrast coding choices in a decade of mixed models. *Journal of Memory and Language*, *125*, 104334. https://doi.org/10.1016/j.jml.2022.104334

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

Brod, G., Lindenberger, U., & Shing, Y. L. (2016). Neural activation patterns during retrieval of schema-related memories: differences and commonalities between children and adults. *Developmental Science*, *20*(6). https://doi.org/10.1111/desc.12475

Brod, G., Lindenberger, U., Werkle-Bergner, M., & Shing, Y. L. (2015). Differences in the neural signature of remembering schema-congruent and schema-incongruent events. *NeuroImage*, *117*, 358–366. https://doi.org/10.1016/j.neuroimage.2015.05.086

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). *The Brain's Default Network*. *Annals of the New York Academy of Sciences*, *1124*(1), 1–38. https://doi.org/10.1196/annals.1440.011

Burgess, N., & Hitch, G. (2005). Computational models of working memory: putting long-term memory into context. *Trends in Cognitive Sciences*, *9*(11), 535–541. https://doi.org/10.1016/j.tics.2005.09.011

Burgess, N., Maguire, E. A., & O'Keefe, J. (2002). The Human Hippocampus and Spatial and Episodic Memory. *Neuron*, *35*(4), 625–641. https://doi.org/10.1016/s0896-6273(02)00830-9

Buuren, M. van, Kroes, M. C. W., Wagner, I. C., Genzel, L., Morris, R. G. M., & Fernández, G. (2014). Initial Investigation of the Effects of an Experimentally Learned Schema on Spatial Associative Memory in Humans. *The Journal of Neuroscience*, *34*(50), 16662–16670. https://doi.org/10.1523/jneurosci.2365-14.2014

Buzsáki, G., & Tingley, D. (2018). Space and Time: The Hippocampus as a Sequence Generator. *Trends in Cognitive Sciences*, *22*(10), 853–869. https://doi.org/10.1016/j.tics.2018.07.006

Caharel, S., & Rossion, B. (2021). The N170 is Sensitive to Long-term (Personal) Familiarity of a Face Identity. *Neuroscience*, *458*, 244–255. https://doi.org/10.1016/j.neuroscience.2020.12.036

Carlson, T. A., Schrater, P., & He, S. (2003). Patterns of Activity in the Categorical Representations of Objects. *Journal of Cognitive Neuroscience*, *15*(5), 704–717. https://doi.org/10.1162/jocn.2003.15.5.704

Carlson, T., Tovar, D. A., Alink, A., & Kriegeskorte, N. (2013). Representational dynamics of object vision: The first 1000 ms. *Journal of Vision*, *13*(10), 1–1. https://doi.org/10.

1167/13.10.1

Carmo-Blanco, N. D., & Allen, J. J. B. (2019). Neural correlates of cue predictiveness during intentional and incidental associative learning: A time-frequency study. *International Journal of Psychophysiology*, *143*, 80–87. https://doi.org/10.1016/j.ijpsycho.2019.06.010

Chen, J., & Bornstein, A. M. (2024). The causal structure and computational value of narratives. *Trends in Cognitive Sciences*, *28*(8), 769–781. https://doi.org/10.1016/j.tics.2024.04.003

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

Chrastil, E. R., Rice, C., Goncalves, M., Moore, K. N., Wynn, S. C., Stern, C. E., & Nyhus, E. (2022). Theta oscillations support active exploration in human spatial navigation. *NeuroImage*, *262*, 119581. https://doi.org/10.1016/j.neuroimage.2022.119581

Cohn-Sheehy, B. I., Delarazan, A. I., Crivelli-Decker, J. E., Reagh, Z. M., Mundada, N. S., Yonelinas, A. P., Zacks, J. M., & Ranganath, C. (2021). Narratives bridge the divide between distant events in episodic memory. *Memory & Cognition*, *50*(3), 478–494. https://doi.org/10.3758/s13421-021-01178-x

Cohn-Sheehy, B. I., Delarazan, A. I., Reagh, Z. M., Crivelli-Decker, J. E., Kim, K., Barnett, A. J., Zacks, J. M., & Ranganath, C. (2021). The hippocampus constructs narrative memories across distant events. *Current Biology*, *31*(22), 4935–4945.e7. https://doi.org/10.1016/j.cub.2021.09.013

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 201–215. https://doi.org/10.1038/nrn755

Crivelli-Decker, J., Hsieh, L.-T., Clarke, A., & Ranganath, C. (2018). Theta oscillations promote temporal sequence learning. *Neurobiology of Learning and Memory*, *153*, 92–

103. https://doi.org/10.1016/j.nlm.2018.05.001

Damasio, A. R. (1989). Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition*, *33*(1-2), 25–62. https://doi.org/10.1016/0010-0277(89)90005-x

Davachi, L., & DuBrow, S. (2015). How the hippocampus preserves order: the role of prediction and context. *Trends in Cognitive Sciences*, *19*(2), 92–99. https://doi.org/10.1016/j.tics.2014.12.004

Daw, N. D., & Dayan, P. (2014). The algorithmic anatomy of model-based evaluation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1655), 20130478. https://doi.org/10.1098/rstb.2013.0478

Dayan, P. (1993). Improving Generalization for Temporal Difference Learning: The Successor Representation. *Neural Computation*, *5*(4), 613–624. https://doi.org/10.1162/neco.1993.5.4.613

Delorme, A. (2023). EEG is better left alone. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-27528-0

Destrebecqz, A., & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychonomic Bulletin & Review*, *8*(2), 343–350. https://doi.org/10.3758/bf03196171

Deuker, L., Olligs, J., Fell, J., Kranz, T. A., Mormann, F., Montag, C., Reuter, M., Elger, C. E., & Axmacher, N. (2013). Memory Consolidation by Replay of Stimulus-Specific Neural Activity. *The Journal of Neuroscience*, *33*(49), 19373–19383. https://doi.org/10.1523/jneurosci.0414-13.2013

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

Dobs, K., Isik, L., Pantazis, D., & Kanwisher, N. (2019). How face perception unfolds over time. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-09239-1

Duff, M. C., Covington, N. V., Hilverman, C., & Cohen, N. J. (2020). Semantic memory and the hippocampus: Revisiting, reaffirming, and extending the reach of their critical relationship. *Frontiers in Human Neuroscience*, *13*. https://doi.org/10.3389/fnhum.2019.00471

Durrant, S. J., Taylor, C., Cairney, S., & Lewis, P. A. (2011). Sleep-dependent consolidation of statistical learning. *Neuropsychologia*, *49*(5), 1322–1331. https://doi.org/10.1016/j.neuropsychologia.2011.02.015

Eichenbaum, H. (2013). Memory on time. *Trends in Cognitive Sciences*, *17*(2), 81–88. https://doi.org/10.1016/j.tics.2012.12.007

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/bf03193146

Fell, J., & Axmacher, N. (2011). The role of phase synchronization in memory processes. *Nature Reviews Neuroscience*, *12*(2), 105–118. https://doi.org/10.1038/nrn2979

Fell, J., Klaver, P., Lehnertz, K., Grunwald, T., Schaller, C., Elger, C. E., & Fernández, G. (2001). Human memory formation is accompanied by rhinal–hippocampal coupling and decoupling. *Nature Neuroscience*, *4*(12), 1259–1264. https://doi.org/10.1038/nn759

Fiebelkorn, I. C., & Kastner, S. (2019). A Rhythmic Theory of Attention. *Trends in Cognitive Sciences*, *23*(2), 87–101. https://doi.org/10.1016/j.tics.2018.11.009

Fortin, N. J., Agster, K. L., & Eichenbaum, H. B. (2002). Critical role of the hippocampus in memory for sequences of events. *Nature Neuroscience*, *5*(5), 458–462. https://doi.org/10.1038/nn834

Foxe, J., & Simpson, G. (2002). Flow of activation from V1 to frontal cortex in humans. *Experimental Brain Research*, *142*(1), 139–150. https://doi.org/10.1007/s00221-001-0906-7

Frank, D., Montaldi, D., Wittmann, B., & Talmi, D. (2018). Beneficial and detrimental effects of schema incongruence on memory for contextual events. *Learning & Memory*, *25*(8), 352–360. https://doi.org/10.1101/lm.047738.118

Frank, D., Montemurro, M. A., & Montaldi, D. (2020). Pattern Separation Underpins Expectation-Modulated Memory. *The Journal of Neuroscience*, *40*(17), 3455–3464. https://doi.org/10.1523/jneurosci.2047-19.2020

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622

Frost, R., Armstrong, B. C., & Christiansen, M. H. (2019). Statistical learning research: A critical review and possible new directions. *Psychological Bulletin*, *145*(12), 1128–1153. https://doi.org/10.1037/bul0000210

Fuentemilla, L., Penny, W. D., Cashdollar, N., Bunzeck, N., & Düzel, E. (2010). Theta-Coupled Periodic Replay in Working Memory. *Current Biology*, *20*(7), 606–612. https://doi.org/10.1016/j.cub.2010.01.057

Garrido, M. I., Barnes, G. R., Kumaran, D., Maguire, E. A., & Dolan, R. J. (2015). Ventromedial prefrontal cortex drives hippocampal theta oscillations induced by mismatch computations. *NeuroImage*, *120*, 362–370. https://doi.org/10.1016/j.neuroimage.2015.07.016

Gauld, A., & Stephenson, G. M. (1967). Some Experiments Relating to Bartlett's Theory of Remembering. *British Journal of Psychology*, *58*(1-2), 39–49. https://doi.org/10.1111/j.2044-8295.1967.tb01054.x

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

Ghosh, V. E., Moscovitch, M., Melo Colella, B., & Gilboa, A. (2014). Schema Representation in Patients with Ventromedial PFC Lesions. *The Journal of Neuroscience*, *34*(36), 12057–12070. https://doi.org/10.1523/jneurosci.0740-14.2014

Gilboa, A., & Marlatte, H. (2017). Neurobiology of Schemas and Schema-Mediated Memory. *Trends in Cognitive Sciences*, *21*(8), 618–631. https://doi.org/10.1016/j.tics.2017.04.013

Gilboa, A., & Moscovitch, M. (2017). Ventromedial prefrontal cortex generates pre-stimulus theta coherence desynchronization: A schema instantiation hypothesis. *Cortex*, *87*, 16–30. https://doi.org/10.1016/j.cortex.2016.10.008

Giuliano, A. E., Bonasia, K., Ghosh, V. E., Moscovitch, M., & Gilboa, A. (2021). Differential Influence of Ventromedial Prefrontal Cortex Lesions on Neural Representations of Schema and Semantic Category Knowledge. *Journal of Cognitive Neuroscience*, 1–28. https://doi.org/10.1162/jocn_a_01746

Gramfort, A. (2013). MEG and EEG data analysis with MNE-python. *Frontiers in Neuroscience*, *7*. https://doi.org/10.3389/fnins.2013.00267

Greve, A., Cooper, E., Kaula, A., Anderson, M. C., & Henson, R. (2017). Does prediction error drive one-shot declarative learning? *Journal of Memory and Language*, *94*, 149–165. https://doi.org/10.1016/j.jml.2016.11.001

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. https://doi.org/10.1037/xge0000498

Griffiths, B. J., Parish, G., Roux, F., Michelmann, S., Plas, M. van der, 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

Grootswagers, T., Wardle, S. G., & Carlson, T. A. (2017). Decoding Dynamic Brain Patterns from Evoked Responses: A Tutorial on Multivariate Pattern Analysis Applied to Time Series Neuroimaging Data. *Journal of Cognitive Neuroscience*, *29*(4), 677–697.

https://doi.org/10.1162/jocn_a_01068

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

Harris, C. R., Millman, K. J., Walt, S. J. van der, Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., Kerkwijk, M. H. van, Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, *585*(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Hastie, R., & Kumar, P. A. (1979). Person memory: Personality traits as organizing principles in memory for behaviors. *Journal of Personality and Social Psychology*, *37*(1), 25–38. https://doi.org/10.1037/0022-3514.37.1.25

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and Overlapping Representations of Faces and Objects in Ventral Temporal Cortex. *Science*, *293*(5539), 2425–2430. https://doi.org/10.1126/science.1063736

Henson, R. N. A. (1998). Short-Term Memory for Serial Order: The Start-End Model. *Cognitive Psychology*, *36*(2), 73–137. https://doi.org/10.1006/cogp.1998.0685

Henson, R. N. A., Norris, D. G., Page, M. P. A., & Baddeley, A. D. (1996). Unchained Memory: Error Patterns Rule out Chaining Models of Immediate Serial Recall. *The Quarterly Journal of Experimental Psychology Section A*, *49*(1), 80–115. https://doi.org/10.1080/713755612

Henson, R. N., & Gagnepain, P. (2010). Predictive, interactive multiple memory systems. *Hippocampus*, *20*(11), 1315–1326. https://doi.org/10.1002/hipo.20857

Herweg, N. A., Solomon, E. A., & Kahana, M. J. (2020). Theta Oscillations in Human

Memory. *Trends in Cognitive Sciences*, *24*(3), 208–227. https://doi.org/10.1016/j.tics.2019.12.006

Heslenfeld, D. J. (2003). *Visual Mismatch Negativity* (pp. 41–59). Springer US. https://doi.org/10.1007/978-1-4615-0294-4_3

Hetenyi, D., Haarsma, J., & Kok, P. (2024). *Pre-stimulus alpha oscillations encode stimulus-specific visual predictions*. http://dx.doi.org/10.1101/2024.03.13.584593

Hitch, G. J., Flude, B., & Burgess, N. (2009). Slave to the rhythm: Experimental tests of a model for verbal short-term memory and long-term sequence learning. *Journal of Memory and Language*, *61*(1), 97–111. https://doi.org/10.1016/j.jml.2009.02.004

Honey, R. C., Watt, A., & Good, M. (1998). Hippocampal Lesions Disrupt an Associative Mismatch Process. *The Journal of Neuroscience*, *18*(6), 2226–2230. https://doi.org/10.1523/jneurosci.18-06-02226.1998

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). https://doi.org/10.1038/ncomms8462

Howard, M. W., & Eichenbaum, H. (2013). The hippocampus, time, and memory across scales. *Journal of Experimental Psychology: General*, *142*(4), 1211–1230. https://doi.org/10.1037/a0033621

Howard, M. W., & Kahana, M. J. (2002). A Distributed Representation of Temporal Context. *Journal of Mathematical Psychology*, *46*(3), 269–299. https://doi.org/10.1006/jmps.2001.1388

Hsieh, L. T., Ekstrom, A. D., & Ranganath, C. (2011). Neural Oscillations Associated with Item and Temporal Order Maintenance in Working Memory. *Journal of Neuroscience*, *31*(30), 10803–10810. https://doi.org/10.1523/jneurosci.0828-11.2011

Hsieh, L.-T., Gruber, Matthias J., Jenkins, Lucas J., & Ranganath, C. (2014). Hippocampal Activity Patterns Carry Information about Objects in Temporal Context. *Neuron*, *81*(5), 1165–1178. https://doi.org/10.1016/j.neuron.2014.01.015

Huang, J., Velarde, I., Ma, W. J., & Baldassano, C. (2023). Schema-based predictive eye movements support sequential memory encoding. *eLife*, *12*. https://doi.org/10.7554/elife.82599

Itier, R. J. (2004). N170 or N1? Spatiotemporal Differences between Object and Face Processing Using ERPs. *Cerebral Cortex*, *14*(2), 132–142. https://doi.org/10.1093/cercor/bhg111

Jackson, R. L. (2021). The neural correlates of semantic control revisited. *NeuroImage*, *224*, 117444. https://doi.org/10.1016/j.neuroimage.2020.117444

Jafarpour, A., Horner, A. J., Fuentemilla, L., Penny, W. D., & Duzel, E. (2013). Decoding oscillatory representations and mechanisms in memory. *Neuropsychologia*, *51*(4), 772–780. https://doi.org/10.1016/j.neuropsychologia.2012.04.002

Jensen, O., & Lisman, J. E. (2005). Hippocampal sequence-encoding driven by a cortical multi-item working memory buffer. *Trends in Neurosciences*, *28*(2), 67–72. https://doi.org/10.1016/j.tins.2004.12.001

Joensen, B. H., Bush, D., Vivekananda, U., Horner, A. J., Bisby, J. A., Diehl, B., Miserocchi, A., McEvoy, A. W., Walker, M. C., & Burgess, N. (2023). Hippocampal theta activity during encoding promotes subsequent associative memory in humans. *Cerebral Cortex*, *33*(13), 8792–8802. https://doi.org/10.1093/cercor/bhad162

Johnston, P., Robinson, J., Kokkinakis, A., Ridgeway, S., Simpson, M., Johnson, S., Kaufman, J., & Young, A. W. (2017). Temporal and spatial localization of prediction-error signals in the visual brain. *Biological Psychology*, *125*, 45–57. https://doi.org/10.1016/j.biopsycho.2017.02.004

Kafkas, A., & Montaldi, D. (2018). Expectation affects learning and modulates memory experience at retrieval. *Cognition*, *180*, 123–134. https://doi.org/10.1016/j.cognition.2018.07.010

Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition*, *24*(1), 103–109. https://doi.org/10.3758/bf03197276

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. *The Journal of Neuroscience*, *17*(11), 4302–4311. https://doi.org/10.1523/jneurosci.17-11-04302.1997

Keller, A. S., Payne, L., & Sekuler, R. (2017). Characterizing the roles of alpha and theta oscillations in multisensory attention. *Neuropsychologia*, *99*, 48–63. https://doi.org/10.1016/j.neuropsychologia.2017.02.021

Kerrén, C., Linde-Domingo, J., Hanslmayr, S., & Wimber, M. (2018). An Optimal Oscillatory Phase for Pattern Reactivation during Memory Retrieval. *Current Biology*, *28*(21), 3383–3392.e6. https://doi.org/10.1016/j.cub.2018.08.065

Kim, H. (2010). Dissociating the roles of the default-mode, dorsal, and ventral networks in episodic memory retrieval. *NeuroImage*, *50*(4), 1648–1657. https://doi.org/10.1016/j.neuroimage.2010.01.051

Kimura, M., Katayama, J., Ohira, H., & Schröger, E. (2009). Visual mismatch negativity: New evidence from the equiprobable paradigm. *Psychophysiology*, *46*(2), 402–409. https://doi.org/10.1111/j.1469-8986.2008.00767.x

King, J.-R., & Dehaene, S. (2014). Characterizing the dynamics of mental representations: the temporal generalization method. *Trends in Cognitive Sciences*, *18*(4), 203–210. https://doi.org/10.1016/j.tics.2014.01.002

Kok, P., Mostert, P., & Lange, F. P. de. (2017). Prior expectations induce prestimulus sensory templates. *Proceedings of the National Academy of Sciences*, *114*(39), 10473–10478. https://doi.org/10.1073/pnas.1705652114

Kolibius, L. D., Roux, F., Parish, G., Ter Wal, M., Van Der Plas, M., Chelvarajah, R., Sawlani, V., Rollings, D. T., Lang, J. D., Gollwitzer, S., Walther, K., Hopfengärtner, R., Kreiselmeyer, G., Hamer, H., Staresina, B. P., Wimber, M., Bowman, H., & Hanslmayr, S. (2023). Hippocampal neurons code individual episodic memories in humans. *Nature Human Behaviour*, *7*(11), 1968–1979. https://doi.org/10.1038/s41562-023-01706-6

Kota, S., Rugg, M. D., & Lega, B. C. (2020). Hippocampal Theta Oscillations Support

Successful Associative Memory Formation. *The Journal of Neuroscience*, *40*(49), 9507–9518. https://doi.org/10.1523/jneurosci.0767-20.2020

Kwon, H., Kronemer, S. I., Christison-Lagay, K. L., Khalaf, A., Li, J., Ding, J. Z., Freedman, N. C., & Blumenfeld, H. (2021). Early cortical signals in visual stimulus detection. *NeuroImage*, *244*, 118608. https://doi.org/10.1016/j.neuroimage.2021.118608

Lad, M., Mullally, S. L., Houston, A. L., Kelly, T., & Griffiths, T. D. (2019). Characterizing memory loss in patients with autoimmune limbic encephalitis hippocampal lesions. *Hippocampus*, *29*(11), 1114–1120. https://doi.org/10.1002/hipo.23150

Lamme, V. A. F., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, *23*(11), 571–579. https://doi.org/10.1016/s0166-2236(00)01657-x

Le Pelley, M. E., Mitchell, C. J., Beesley, T., George, D. N., & Wills, A. J. (2016). Attention and associative learning in humans: An integrative review. *Psychological Bulletin*, *142*(10), 1111–1140. https://doi.org/10.1037/bul0000064

Lee, H., & Chen, J. (2022). Predicting memory from the network structure of naturalistic events. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-022-31965-2

Lega, B., Burke, J., Jacobs, J., & Kahana, M. J. (2014). Slow-Theta-to-Gamma Phase–Amplitude Coupling in Human Hippocampus Supports the Formation of New Episodic Memories. *Cerebral Cortex*, *26*(1), 268–278. https://doi.org/10.1093/cercor/bhu232

Lin, J.-J., Rugg, M. D., Das, S., Stein, J., Rizzuto, D. S., Kahana, M. J., & Lega, B. C. (2017). Theta band power increases in the posterior hippocampus predict successful episodic memory encoding in humans. *Hippocampus*, *27*(10), 1040–1053. https://doi.org/10.1002/hipo.22751

Lisman, J. E., Talamini, L. M., & Raffone, A. (2005). Recall of memory sequences by interaction of the dentate and CA3: A revised model of the phase precession. *Neural Networks*, *18*(9), 1191–1201. https://doi.org/10.1016/j.neunet.2005.08.008

Lisman, J., & Redish, A. D. (2009). Prediction, sequences and the hippocampus. *Philo-*

sophical Transactions of the Royal Society B: Biological Sciences, *364*(1521), 1193–1201. https://doi.org/10.1098/rstb.2008.0316

Liu, Y., Nour, M. M., Schuck, N. W., Behrens, T. E. J., & Dolan, R. J. (2022). Decoding cognition from spontaneous neural activity. *Nature Reviews Neuroscience*, *23*(4), 204–214. https://doi.org/10.1038/s41583-022-00570-z

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*(4), 1122–1135. https://doi.org/10.3758/s13428-014-0532-5

Maguire, E. A., Burgess, N., & O'Keefe, J. (1999). Human spatial navigation: cognitive maps, sexual dimorphism, and neural substrates. *Current Opinion in Neurobiology*, *9*(2), 171–177. https://doi.org/10.1016/s0959-4388(99)80023-3

Mandler, J. M., & Johnson, N. S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, *9*(1), 111–151. https://doi.org/10.1016/0010-0285(77)90006-8

Mäntylä, T., & Bäckman, L. (1992). Aging and memory for expected and unexpected objects in real-world settings. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(6), 1298–1309. https://doi.org/10.1037/0278-7393.18.6.1298

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

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information.* https://doi.org/10.7551/mitpress/9780262514620.001.0001

Mayes, A. R., Holdstock, J. S., Isaac, C. L., Montaldi, D., Grigor, J., Gummer, A., Cariga, P., Downes, J. J., Tsivilis, D., Gaffan, D., Gong, Q., & Norman, K. A. (2004). Associative recognition in a patient with selective hippocampal lesions and relatively normal item recognition. *Hippocampus*, *14*(6), 763–784. https://doi.org/10.1002/hipo.10211

McClelland, J. L. (2013). Incorporating rapid neocortical learning of new schema-consistent information into complementary learning systems theory. *Journal of Experimental Psychology: General*, *142*(4), 1190–1210. https://doi.org/10.1037/a0033812

McDermott, T. J., Wiesman, A. I., Proskovec, A. L., Heinrichs-Graham, E., & Wilson, T. W. (2017). Spatiotemporal oscillatory dynamics of visual selective attention during a flanker task. *NeuroImage*, *156*, 277–285. https://doi.org/10.1016/j.neuroimage.2017.05.014

Millidge, B., Seth, A., & Buckley, C. L. (2021). *Predictive coding: A theoretical and experimental review*. https://doi.org/10.48550/ARXIV.2107.12979

Moccia, A., Plummer, M., Simpson, I., & Morcom, A. M. (2022). *The effects of external cue overlap and internal goals on selective memory retrieval as revealed by electroencephalographic (EEG) neural pattern reinstatement*. http://dx.doi.org/10.1101/2022.10.21.513221

Morrison, A. B., Conway, A. R. A., & Chein, J. M. (2014). Primacy and recency effects as indices of the focus of attention. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00006

Moscovitch, M. (1997). Strategic retrieval and the frontal lobes: Evidence from confabulation and amnesia. *Neuropsychologia*, *35*(7), 1017–1034. https://doi.org/10.1016/s0028-3932(97)00028-6

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

Nadel, L., Winocur, G., Ryan, L., & Moscovitch, M. (2007). Systems consolidation and hippocampus: two views. *Debates in Neuroscience*, *1*(2-4), 55–66. https://doi.org/10.1007/s11559-007-9003-9

Nakazawa, K., Quirk, M. C., Chitwood, R. A., Watanabe, M., Yeckel, M. F., Sun, L. D., Kato, A., Carr, C. A., Johnston, D., Wilson, M. A., & Tonegawa, S. (2002). Requirement for Hippocampal CA3 NMDA Receptors in Associative Memory Recall. *Science*, *297*(5579), 211–218. https://doi.org/10.1126/science.1071795

Nolden, S., Turan, G., Bein, O., Davachi, L., & Shing, Y. L. (2025). The impact of mnemonic prediction errors on episodic memory: A lifespan study. *Developmental Psychology*. https://doi.org/10.1037/dev0001966

O'Kane, G., Kensinger, E. A., & Corkin, S. (2004). Evidence for semantic learning in profound amnesia: An investigation with patient H.M. *Hippocampus*, *14*(4), 417–425. https://doi.org/10.1002/hipo.20005

O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, *34*(1), 171–175. https://doi.org/10.1016/0006-8993(71)90358-1

Oberauer, K. (2003). Understanding serial position curves in short-term recognition and recall. *Journal of Memory and Language*, *49*(4), 469–483. https://doi.org/10.1016/s0749-596x(03)00080-9

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

Ortiz-Tudela, J., Turan, G., Vilas, M., Melloni, L., & Shing, Y. L. (2024). Schema-driven prediction effects on episodic memory across the lifespan. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *379*(1913). https://doi.org/10.1098/rstb.2023.0401

Page, M. P. A., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review*, *105*(4), 761–781. https://doi.org/10.1037/0033-295x.105.4.761-781

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, *87*(6), 532–552. https://doi.org/10.1037/0033-295x.87.6.532

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in python. *J. Mach. Learn. Res.*, *12*(null), 28252830.

Preston, Alison R., & Eichenbaum, H. (2013). Interplay of Hippocampus and Prefrontal Cortex in Memory. *Current Biology*, *23*(17), R764–R773. https://doi.org/10.1016/j.cub.2013.05.041

Quent, J. A., Greve, A., & Henson, R. N. (2022). Shape of U: The Nonmonotonic Relationship Between Object–Location Memory and Expectedness. *Psychological Science*, *33*(12), 2084–2097. https://doi.org/10.1177/09567976221109134

Quent, J. A., Henson, R. N., & Greve, A. (2021). A predictive account of how novelty influences declarative memory. *Neurobiology of Learning and Memory*, *179*, 107382. https://doi.org/10.1016/j.nlm.2021.107382

Quiroga, R. Q. (2012). Concept cells: the building blocks of declarative memory functions. *Nature Reviews Neuroscience*, *13*(8), 587–597. https://doi.org/10.1038/nrn3251

Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C., & Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature*, *435*(7045), 1102–1107. https://doi.org/10.1038/nature03687

R Core Team. (2023). *R: A Language and Environment for Statistical computing.* R Foundation for Statistical Computing. https://www.R-project.org/

Ralph, M. A. L., Jefferies, E., Patterson, K., & Rogers, T. T. (2016). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, *18*(1), 42–55. https://doi.org/10.1038/nrn.2016.150

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87. https://doi.org/10.1038/4580

Raykov, P. P., Keidel, J. L., Oakhill, J., & Bird, C. M. (2019). The brain regions supporting schema-related processing of people's identities. *Cognitive Neuropsychology*, *37*(1-2), 8–24. https://doi.org/10.1080/02643294.2019.1685958

Recasens, M., Gross, J., & Uhlhaas, P. J. (2018). Low-Frequency Oscillatory Correlates of Auditory Predictive Processing in Cortical-Subcortical Networks: A MEG-Study. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-32385-3

Reddy, L., Self, M. W., Zoefel, B., Poncet, M., Possel, J. K., Peters, J. C., Baayen, J. C., Idema, S., VanRullen, R., & Roelfsema, P. R. (2021). Theta-phase dependent neuronal coding during sequence learning in human single neurons. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-25150-0

Rescorla, R., & Wagner, A. (1972). A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In *Classical Conditioning II: Current Research and Theory: Vols. Vol. 2* (pp. 64–99).

Robin, J., & Moscovitch, M. (2017). Details, gist and schema: hippocampal–neocortical interactions underlying recent and remote episodic and spatial memory. *Current Opinion in Behavioral Sciences*, *17*, 114–123. https://doi.org/10.1016/j.cobeha.2017.07.016

Robinson, J. E., Breakspear, M., Young, A. W., & Johnston, P. J. (2018). Dose-dependent modulation of the visually evoked N1/N170 by perceptual surprise: a clear demonstration of prediction-error signalling. *European Journal of Neuroscience*, *52*(11), 4442–4452. https://doi.org/10.1111/ejn.13920

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

Rosenbaum, R. S., Köhler, S., Schacter, D. L., Moscovitch, M., Westmacott, R., Black, S. E., Gao, F., & Tulving, E. (2005). The case of K.C.: contributions of a memory-impaired person to memory theory. *Neuropsychologia*, *43*(7), 989–1021. https://doi.org/10.1016/j.neuropsychologia.2004.10.007

Russo, N., Hagmann, C. E., Andrews, R., Black, C., Silberman, M., & Shea, N. (2018). Validation of the C.A.R.E. stimulus set of 640 animal pictures: Name agreement and quality ratings. *PLOS ONE*, *13*(2), e0192906. https://doi.org/10.1371/journal.pone.0192906

Schacter, D. L., Moscovitch, M., Tulving, E., McLachlan, D. R., & Freedman, M. (1986). Mnemonic precedence in amnesic patients: An analogue of the AB error in infants? *Child Development*, *57*(3), 816. https://doi.org/10.2307/1130357

Schapiro, A., Kustner, Lauren V., & Turk-Browne, Nicholas B. (2012). Shaping of Object Representations in the Human Medial Temporal Lobe Based on Temporal Regularities. *Current Biology*, *22*(17), 1622–1627. https://doi.org/10.1016/j.cub.2012.06.056

Schapiro, A., & Turk-Browne, N. (2015). *Statistical Learning* (pp. 501–506). Elsevier. https://doi.org/10.1016/b978-0-12-397025-1.00276-1

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery & Psychiatry*, *20*(1), 11–21. https://doi.org/10.1136/jnnp.20.1.11

Sergent, J., Ohta, S., & Macdonald, B. (1992). Functional neuroanatomy of face and object processing. *Brain*, *115*(1), 15–36. https://doi.org/10.1093/brain/115.1.15

Sherman, B. E., Graves, K. N., Huberdeau, D. M., Quraishi, I. H., Damisah, E. C., & Turk-Browne, N. B. (2022). Temporal Dynamics of Competition between Statistical Learning and Episodic Memory in Intracranial Recordings of Human Visual Cortex. *Journal of Neuroscience*, *42*(48), 9053–9068. https://doi.org/10.1523/JNEUROSCI.0708-22.2022

Sherman, B. E., & Turk-Browne, N. B. (2020). Statistical prediction of the future impairs episodic encoding of the present. *Proceedings of the National Academy of Sciences*, *117*(37), 22760–22770. https://doi.org/10.1073/pnas.2013291117

Sinclair, A. H., Manalili, G. M., Brunec, I. K., Adcock, R. A., & Barense, M. D. (2021). Prediction errors disrupt hippocampal representations and update episodic memories. *Proceedings of the National Academy of Sciences*, *118*(51). https://doi.org/10.1073/

pnas.2117625118

Smith, M. E., & Loschky, L. C. (2019). The influence of sequential predictions on scene-gist recognition. *Journal of Vision*, *19*(12), 14. https://doi.org/10.1167/19.12.14

Smith, M., Fitzgibbons, C., Faiola, A., & Loschky, L. (2021). Rapid scene categorization is not purely feed-forward: An EEG investigation of scene gist facilitation by sequential predictions. *Journal of Vision*, *21*(9), 2898. https://doi.org/10.1167/jov.21.9.2898

Spalding, K. N., Jones, S. H., Duff, M. C., Tranel, D., & Warren, D. E. (2015). Investigating the Neural Correlates of Schemas: Ventromedial Prefrontal Cortex Is Necessary for Normal Schematic Influence on Memory. *The Journal of Neuroscience*, *35*(47), 15746–15751. https://doi.org/10.1523/jneurosci.2767-15.2015

Stachenfeld, K. L., Botvinick, M. M., & Gershman, S. J. (2017). The hippocampus as a predictive map. *Nature Neuroscience*, *20*(11), 1643–1653. https://doi.org/10.1038/nn.4650

Staresina, B. P., Alink, A., Kriegeskorte, N., & Henson, R. N. (2013). Awake reactivation predicts memory in humans. *Proceedings of the National Academy of Sciences*, *110*(52), 21159–21164. https://doi.org/10.1073/pnas.1311989110

Staresina, B. P., Gray, J. C., & Davachi, L. (2008). 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

Staresina, B. P., & Wimber, M. (2019). A Neural Chronometry of Memory Recall. *Trends in Cognitive Sciences*, *23*(12), 1071–1085. https://doi.org/10.1016/j.tics.2019.09.011

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

Stefanics, G., KremlÃ¡Ä ek, J., & Czigler, I. (2014). Visual mismatch negativity: A predictive coding view. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00666

Takehara-Nishiuchi, K. (2020). Prefrontal–hippocampal interaction during the encoding of new memories. *Brain and Neuroscience Advances*, *4*, 239821282092558. https://doi.org/10.1177/2398212820925580

Tarder-Stoll, H., Baldassano, C., & Aly, M. (2024). Consolidation Enhances Sequential Multistep Anticipation but Diminishes Access to Perceptual Features. *Psychological Science*, *35*(10), 1178–1199. https://doi.org/10.1177/09567976241256617

Teyler, T. J., & Rudy, J. W. (2007). The hippocampal indexing theory and episodic memory: Updating the index. *Hippocampus*, *17*(12), 1158–1169. https://doi.org/10.1002/hipo.20350

Thakral, P. P., Wang, T. H., & Rugg, M. D. (2017). Decoding the content of recollection within the core recollection network and beyond. *Cortex*, *91*, 101–113. https://doi.org/10.1016/j.cortex.2016.12.011

Treder, M. S. (2020). MVPA-light: A classification and regression toolbox for multi-dimensional data. *Frontiers in Neuroscience*, *14*. https://doi.org/10.3389/fnins.2020.00289

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

Tulving, E. (2002). Episodic Memory: From Mind to Brain. *Annual Review of Psychology*, *53*(1), 1–25. https://doi.org/10.1146/annurev.psych.53.100901.135114

Tulving, E., & Kroll, N. (1995). Novelty assessment in the brain and long-term memory encoding. *Psychonomic Bulletin & Review*, *2*(3), 387–390. https://doi.org/10.3758/bf03210977

van Kesteren, M. T. R., Beul, S. F., Takashima, A., Henson, R. N., Ruiter, D. J., & Fernández, G. (2013). Differential roles for medial prefrontal and medial temporal cortices in schema-dependent encoding: From congruent to incongruent. *Neuropsychologia*, *51*(12), 2352–2359. https://doi.org/10.1016/j.neuropsychologia.2013.05.027

van Kesteren, M. T. R., Fernández, G., Norris, D. G., & Hermans, E. J. (2010). Persistent schema-dependent hippocampal-neocortical connectivity during memory encoding and postencoding rest in humans. *Proceedings of the National Academy of Sciences*, *107*(16), 7550–7555. https://doi.org/10.1073/pnas.0914892107

van Kesteren, M. T. R., & Meeter, M. (2020). How to optimize knowledge construction in the brain. *Npj Science of Learning*, *5*(1). https://doi.org/10.1038/s41539-020-0064-y

van Kesteren, M. T. R., Ruiter, D. J., Fernández, G., & Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends in Neurosciences*, *35*(4), 211–219. https://doi.org/10.1016/j.tins.2012.02.001

Voeten, C. C. (2020). *Buildmer: Stepwise elimination and term reordering for mixed-effects regression.* https://CRAN.R-project.org/package=buildmer.

Warren, D. E., Jones, S. H., Duff, M. C., & Tranel, D. (2014). False Recall Is Reduced by Damage to the Ventromedial Prefrontal Cortex: Implications for Understanding the Neural Correlates of Schematic Memory. *The Journal of Neuroscience*, *34*(22), 7677–7682. https://doi.org/10.1523/jneurosci.0119-14.2014

Wheeler, M. E., Petersen, S. E., & Buckner, R. L. (2000). Memory's echo: Vivid remembering reactivates sensory-specific cortex. *Proceedings of the National Academy of Sciences*, *97*(20), 11125–11129. https://doi.org/10.1073/pnas.97.20.11125

Yacoby, A., Greve, A., Schaverien, E., Shtoots, L., Henson, R., & Levy, D. (2025). *The Effects of Prior Knowledge on Episodic Memory Depend on Retrieval Goals and Orientation.* https://doi.org/10.23668/PSYCHARCHIVES.16493