Exploring how the presence of a pattern affects both memory and generalisation

Jamie Philip Cockcroft

Doctor of Philosophy

University of York

Psychology

November 2021

Abstract

Many theories of schema-based processing implicitly assume that information irrelevant to a schema will be unaffected by its presence. However, this notion has yet to be formally examined. In the present work, the precision paradigm was used to explore memory-based generalisation as a means of assessing schematic influence for both relevant and irrelevant information. Here, participants learned word-location associations, with one group of words having locations grouped in one area of the circle (clustered condition) and the other having locations distributed uniformly (non-clustered condition). Using a series of approaches (i.e., behavioural, computational modelling, and neuroimaging), the present thesis explored how schematic information influenced behaviour. It was found that the presence of a pattern does impact information that could be considered irrelevant to the pattern itself. This result suggests that the use of schema-irrelevant controls, or theories ignoring possible biases produced by schematic information, may need to be updated. However, alternative explanations for these effects were proposed and examined (e.g., interference) using computational modelling. Here it was found that proximity- (i.e., items close together will have a reduced probability of retrieval) and semantic-based (i.e., spreading of activation for similar items in memory) interference could produce these patterns of effects. This opens up further questions regarding whether other studies implementing the same paradigm have studied schema or another process entirely. Finally, the neural underpinnings of memory-based generalisation were explored via a preliminary analysis of pilot data. These results suggested that the ventromedial prefrontal cortex and dorsal striatum may play important roles in memory. However, some design and analysis considerations were proposed to assess these effects more closely. Overall, the present thesis provides clear evidence that the presence of a pattern can affect both memory and generalisation for both relevant and irrelevant information.

Keywords*: memory, generalisation, schema, interference*

List of Contents

List of Tables

List of Figures

Accompanying Material

The material presented in Chapter 2 and the analyses reported in Chapter 3 have an associated page on the Open Science Framework, containing pre-registrations, experimental materials, data and analyses scripts. These can be located here: https://osf.io/bxru4/.

Acknowledgements

First and foremost, a special thanks goes to Dr Aidan Horner. Back in 2017 when I first approached you about doing a PhD, you were enthusiastic about the project and eager to dive in. This enthusiasm has never wavered and you continuously encouraged me throughout the PhD. You have supported me with getting funding and mentored me throughout the entirety of the PhD and, as a result, made me a better scientist. You made me think critically about experimental design and analysis, theory application, and allowed me the opportunity to develop a broader range of skills than I had imagined I would during the PhD (e.g., computational modelling). You also supported me in all of my additional work (e.g., YLTA, student representative, ReproducibiliTea), always offering helpful advice and guiding me along the way when I began to struggle. You have been a fantastic mentor who I am beyond appreciative to have had support me throughout this journey and cannot believe you managed to put up with my antics for the entirety of my time here. I have been extremely fortunate.

Next, we have Professor Gareth Gaskell. You have challenged me throughout the PhD to think more broadly, to consider alternative interpretations and to avoid anthropomorphising models (though this latter one is more challenging than you would think!) You have been a great secondary supervisor and always manage to keep me on my toes, which has been great in enhancing my ability to deal with challenging questions on the spot and my experimental design skills. Due to your challenges, I have grown to identify how different areas may help with interpreting unusual findings and how to deal with questions that I would not have been able to in my first year. Thank you for your support.

Another huge thanks go to Dr Sam Berens. I came to this PhD with the belief that I understood statistics. However, you showed me what it means to understand statistics. You provided me with mentorship to learn the complexities of statistical modelling, Bayesian statistics and were integral in helping me understand some of the nuance surrounding why dummy-coding does not provide you with the "main effects". You also helped support my development in working with neuroimaging datasets and have been a great help throughout this PhD journey – so thank you!

Further thanks go to my Thesis Advisory Panel members: Dr Lisa Henderson and Professor Elizabeth Jefferies. Both of you provided me with an opportunity to develop my presenting skills, address challenging, though theoretically interesting, questions and were always supportive of me throughout the entire PhD. You also provided me with a safe environment to cover findings that didn't make much sense at the time. Your feedback and patience have been paramount to making this thesis what it is today. So, thank you!

Lab Members past (Dr Bardur Joensen, Dr Jade Pickering, Dr Nora Andermane, Dr Emma James) and present (Dr Jennifer Ashton). You have all supported me and shown me what it means to be a scientist – you all have superbly demonstrated what a true scientist looks like and how to make a real impact in the realms of academia. You also made me jealous with how incredibly great your work is.

Next, I would like to thank friends, both from inside and outside my academic circle. Zoe McBoyle, you have provided me with an opportunity to rant and rave about academia and gave me space to present some of my findings to a lay audience – for that, I am hugely appreciative. Dr Holly Brown, you have helped me throughout the entirety of this PhD. You have been a rock, supported me through a lot of trials and tribulations, but also pushed me to experience new and exciting (sometimes terrifying) things. As a result, I learned what it means to be an early career researcher as you are a shining star among us. Andrea Woodward, you have been an incredible person to meet and you were beyond supportive from day one up until the day you abandoned us to pursue a PhD yourself (not that I am bitter in any way). Thank you for being there and always keeping me on track.

To other friends along the way (i.e., Nick Souter, Anna á Váli Guttesen, Catia Oliveira, Juliana Olivier, Alex Mepham, Dr Alex Benjamin, and Delali Konu) thank you! You have done some awesome work and provided me with support by allowing me a safe space to vent when needed. I hope I have done the same for you.

Of course, further thanks go to my family and my partner, especially my parents Katheren and Philip Cockcroft. They have had to put up with a lot of anxiety resulting from academic stress, future prospects and just me being me. Thank you for always being there and supporting my ambitions.

Finally, a thanks go to the participants who took part in the experiments I conducted – I know the task was a little bit long-winded, but you have hopefully helped us to better understand schematic processes. Thanks go to the Department of Psychology, University of York and the York Neuroimaging Centre, all of whom have provided me with a space to develop as a scientist and funded my projects.

Author's Declaration

This thesis is a presentation of original work completed solely by the author, under the supervision of Dr Aidan J. Horner and Professor M. Gareth Gaskell. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Publications

Elements of the thesis have been submitted for publication in a scientific journal and released as a pre-print. The main chapter this affects is Chapter 2. However, elements of the publication also appear in other chapters. Specifically, Chapter 1, 3 and 5. The reference for the pre-print is:

Cockcroft, J. P., Berens, S., Gaskell, M., & Horner, A. J. (2021, August 24). *Schematic information influences memory and generalisation behaviour for schema-relevant and -irrelevant information*. https://doi.org/10.31234/osf.io/nzurq.

Conference Presentations

Some of the findings presented within the thesis have also been presented at the following conferences:

Chapter 2

Cockcroft, J. P., Berens, S. C., Gaskell, M. G., Horner, A. J. (2021, January). *Schema influence both behaviour for both schema-relevant and -irrelevant information.* Poster session presented at the Experimental Psychology Society Meeting, Online. [PDF]

Cockcroft, J. P., Berens, S. C., Gaskell, M. G., Horner, A. J. (2019, December). *Back to the present: How memory guides decision behaviour.* Conference talk at the Greater Yorkshire Memory Meeting, York, UK. [Slides]

Cockcroft, J. P., Berens, S. C., Gaskell, M. G., Horner, A. J. (2019, October). *Back to the present: How memory guides decision behaviour.* Conference talk at the Learning and Memory Processes Conference, Durham, UK. [Slides]

Chapter 3

Though the largest portion of findings came from Chapter 2, the exploratory and confirmatory analyses reported in Chapter 3 were also presented at EPS:

Cockcroft, J. P., Berens, S. C., Gaskell, M. G., Horner, A. J. (2021, January). *Schema influence both behaviour for both schema-relevant and -irrelevant information.* Poster session presented at the Experimental Psychology Society Meeting, Online. [PDF]

Use of Secondary Data

Within the present thesis, a secondary data analysis was conducted using the data from Berens et al. (2020) with permission granted by the senior author. This was done to support the conclusions of an analysis in Chapter 3. The data file used can be found on the Open Science Framework: https://osf.io/6mx3s/. The full reference for the paper is:

Berens, S. C., Richards, B. A., & Horner, A. J. (2020). Dissociating memory accessibility and precision in forgetting. *Nature Human Behaviour, 4*(8), 866-877. https://doi.org/10.1038/s41562-020-0888-8

.

Chapter 1:
Literature Review – Memory, Schema and Generalisation

Part of the present literature review's content was previously published as a preprint: Cockcroft, J. P., Berens, S., Gaskell, M., & Horner, A. J. (2021, August 24). *Schematic information influences memory and generalisation behaviour for schema-relevant and -irrelevant information*. https://doi.org/10.31234/osf.io/nzurq. Some content was moved to the present Literature Review from the Introduction of the preprint in order to give more context to the thesis.

Evidence shows that predictions about future events are often based on past experience (Preston et al., 2004; Zeithamova et al., 2012). Such predictions require the generalisation of existing knowledge to novel instances. The present review aims to provide an overview of the processes involved in memory-based generalisation and how questions surrounding schema and generalisation are to be explored within the present thesis. To do this, the review considers the episodic-semantic distinction, how memories change over time, schema influence on memory and generalisation and the neural processes involved when generalising to novel instances. The Chapter ends with an overview of the thesis procedures and individual chapter aims.

1.1. Declarative Memory: An Episodic-Semantic Continuum

Memory systems allow us to encode, store and retrieve information (Tulving, 1972, 1986). Several memory systems are hypothesised to exist, with evidence indicating that each system is functionally distinct (Graham et al., 2000; Takashima et al., 2009; Tulving, 2002). The long-term store, the primary focus of the present thesis, has two distinct memory types: declarative (explicit) and non-declarative (implicit). Declarative memories can be brought to conscious awareness, whilst non-declarative influences behaviour without conscious awareness (Squire, 2004). Further distinctions exist, but the focus of the present discussion will be on the declarative system and its stores: episodic and semantic memory (though see McKoon & Ratcliff, 1986 and Toth & Hunt, 1999 for critiques regarding these proposed dissociations).

Tulving (1972) provided a conceptual framework for episodic and semantic memory. Episodic memory refers to the capacity to re-experience a past episode by re-activating the spatio-temporal context and individual elements of an event. In contrast, semantic memory refers to memory for general knowledge about the world. Knowing that Leonardo da Vinci's

"The Last Supper" is in the Santa Maria delle Grazie in Milan would require semantic memory. In contrast, remembering when you visited Santa Maria delle Grazie and saw the painting yourself would require episodic memory.

Evidence supports the view of an episodic-semantic distinction (Greenberg & Verfaellie, 2010). For instance, Vargha-Khadem et al. (1997) found that early bilateral hippocampal atrophy resulted in episodic memory deficits but relatively intact semantic memory. Despite episodic memory loss, the children studied in Vargha-Khadem et al. (1997) developed similar levels of written and verbal comprehension to their peers. In a similar vein, patients with semantic dementia show relatively intact episodic memory, at least in the early stages of the disease (Graham et al., 2000).

Despite evidence suggesting a dissociation, more recent views speculate that the episodic-semantic distinction may be better represented as a continuum (e.g., Greenberg & Verfaellie, 2010; Irish & Vatansever, 2020). Irish and Vatansever (2020) proposed that the episodic-semantic network can be viewed as a gradient from detail-rich (i.e., episodic memory) to abstract, less detailed (i.e., semantic memory) representations. From this perspective, the current goal or context will influence the memory representations that are relied upon. Support for this view of an interplay, or spectrum, can be shown in the Vargha-Khadem et al. (1997) study addressed above. Despite the preserved semantic memory of patients in Vargha-Khadem et al. (1997), there were still some impairments, albeit less pronounced than the episodic deficit. Further, patients with hippocampal lesions, leading to episodic memory impairments, also show slower learning of new information that is often less well integrated with existing knowledge (Greenberg & Verfaellie, 2010).

Similarly, work has shown how semantic memory can influence the recall of episodic information. Bartlett (1932) had participants recall a Native American ghost story at different intervals: within a few hours, days, weeks or even years. During recall, participants often altered aspects of the story to fit with their semantic knowledge (e.g., seal hunting became fishing) and rationalised elements (e.g., removal of spiritual elements). According to Bartlett (1932), the reason for these changes was schema – abstract representations of knowledge that bias the encoding and retrieval of episodic memory. Overall, this demonstrates an interplay between the two systems supporting the view that they may not be entirely distinct but instead interact.

1.1.1. Memory Consolidation: How Memory Representations Change

Schemas represent commonalities across multiple experiences (this is discussed in more detail below, see: Schema) and are believed to develop via systems consolidation (McClelland et al., 1995). Memory consolidation is defined as a post-encoding process of reorganisation, stabilisation, and qualitative changes in memory representation, from concrete (i.e., episodic) to abstract (i.e., semantic; McClelland et al., 1995; Winocur et al., 2010). These more semantic representations are often referred to as schema (Ghosh & Gilboa, 2014).

Memory consolidation has been shown to occur during offline periods, such as sleep or rest (Gais et al., 2006; Hu et al., 2006; Mograss et al., 2009; Wagner et al., 2007), but has also been observed during online (wake) periods, driven by processes such as retrieval practice (Antony et al., 2017; Ferreira et al., 2019). Through consolidation, regularities across experiences are thought to be extracted and utilised for future behaviour (Batterink & Paller, 2017; Mirković et al., 2019; Sweegers & Talamini, 2014). Though sleep-based consolidation is not pertinent to the present thesis, understanding how memory representations change is. Some prominent theories of system-based consolidation are Complementary Learning

Systems (CLS; McClelland et al., 1995), Multiple Trace Theory (MTT; Nadel & Moscovitch, 1997) and Trace Transformation Theory (TTT; Winocur et al., 2007).

1.1.1.1. Complementary Learning Systems (McClelland et al., 1995)

The CLS (McClelland et al., 1995) model argues there are two systems in memory – one is a fast-encoding system (i.e., episodic memory) within the hippocampus, the other is a slower neocortically based system (i.e., semantic memory). The hippocampal system will quickly learn the pattern of neocortical activity present at the time of the event and reinstate this pattern during recall. The neocortical system will slowly learn this pattern of cortical reinstatement, meaning the event can be recalled independently of the hippocampus. According to this model, the hippocampus will facilitate the reinstatement of overlapping memories, which will result in regularities across events being extracted and arbitrary contextual information being forgotten. In other words, episodic memories become semanticised over time.

Support for this view has come from many studies (see McClelland et al., 1995, for review). One compelling area of work comes from evidence of a temporal gradient following hippocampal lesions. A temporal gradient occurs when memories immediately preceding the amnesia onset are not well remembered, whilst those further from this event are better remembered. Research has shown that amnesic patients who experienced bilateral hippocampal damage show a temporal gradient when remembering experiences (Albert et al., 1979; Beatty et al., 1988; Reed & Squire, 1998; Scoville & Milner, 1957; Tulving et al., 1988). In other words, their memory is often better for events further back in time post-amnesia onset. These findings are also supported by work in non-human animals (Maren et al., 1997; Winocur, 1990; Zola-Morgan & Squire, 1990; but see Frankland & Bontempi, 2005, for review). This demonstration of a temporal gradient shows how memories change over time and become independent of the hippocampus, shifting from episodic to semantic and thus supporting the CLS proposals.

1.1.1.2. Multiple Trace Theory (Nadel & Moscovitch, 1997)

Nevertheless, one of the problems with the CLS model is that it cannot account for some of the findings reported within the literature. For instance, some patients have restricted damage to the hippocampus but do not show temporally graded amnesia (Cipolotti et al., 2001; Rosenbaum et al., 2009; Spiers et al., 2001), suggesting that these individuals' memories did not become hippocampally independent. If it is the case that memories always become hippocampally independent, then why is it that some patients show no evidence of a temporal gradient following hippocampal lesions? To address this, MTT (Nadel & Moscovitch, 1997) was developed. MTT argues that instead of episodic memories becoming hippocampally independent, two memory traces co-exist – an episodic (within the hippocampus) and semantic (within the neocortex). When the task requires the contextual and spatial information of the event, then the hippocampus will still be required. However, when such information is not required, the neocortical (more abstract) representation can be activated without the need for the hippocampus. This model fits closely with the arguments made by Irish and Vatansever (2020) that episodic and semantic memory systems can be viewed as part of a continuum, with the task itself determining which type of memory representation is relied upon.

More specifically, MTT proposes that when an event is encoded, information within an event that shares commonalities to other experiences will also be re-activated and undergo further encoding (Moscovitch & Gilboa, 2021). Therefore, each new experience will reactive similar experiences from memory so that the newly encoded context can be associated with these past events. For instance, if a fox terrier was observed being walked in the park as part of the encoded event, then the trace would become associated with other memory traces associated with this breed of dog. Whilst this sole event will be distinguishable from the other events, it will feed into the semantic trace of a fox terrier. As a result, independent episodic traces can exist (e.g., the fox terrier being walked through the park), whilst also being associated with semantic traces (e.g., what a fox terrier looks like). This would fit with findings of complete episodic memory loss following hippocampal damage, but the retention of semantic traces (which are independent of the hippocampus). For instance, the ability to retrieve coarser information related to a spatial layout despite a lack of memory for individual experiences within that layout (Rosenbaum et al., 2000).

1.1.1.3. Trace Transformation Theory (Winocur et al., 2007)

Building on MTT, TTT (Winocur et al., 2007) proposed that whilst episodic memories would rely on the hippocampus when context-specific information was required, it did not mean that the experiences themselves did not change. Specifically, over time, episodic memories would be transformed such that some aspects were emphasised (i.e., those that shared common properties across events) and others de-emphasised (i.e., contextual details that may not be useful for future behaviour). This transformation would be based on existing knowledge and subsequent experience (Robin & Moscovitch, 2017). In other words, the episodic trace would become increasingly gist-like (i.e., lose contextual information, whilst retaining some specifics about the episode) and even schematic (i.e., a loss of any specific episode-unique details, but extracting commonalities across a series of related experiences) in most instances. This does not mean the episodic trace would be completely lost. Similar to MTT, TTT argued that these representations could co-exist and even interact depending on the requirements of a particular task. However, unlike MTT, TTT argues that as episodic memories become less necessary, there will be a gradual loss of these representations over time. Both MTT and TTT are discussed in more detail in Moscovitch and Gilboa (2021).

1.1.2. Summary

Historically, a division between episodic and semantic memory was proposed. Though evidence supports a dissociation between these two systems, it is clear that they interact. According to the CLS (and TTT) model, memories become semanticised over time through consolidation; specific episodic details are lost as the more abstract components of related episodes are extracted. However, evidence does suggest that memories may not always become hippocampally independent. As described by MTT, both an episodic and semantic trace may be present simultaneously, with episodic experiences being continually reliant on the hippocampus. Building on this proposal, TTT suggests that episodic and semantic traces may co-exist, as argued by MTT, but also acknowledges how episodic traces may be forgotten over time as they become less useful.

1.2. Schema

1.2.1. What is (and is not) a Schema?

Schemas are a form of memory representation associated with semantic memory (Radvansky & Tamplin, 2012) and represent the abstraction of commonalities across multiple experiences (Ghosh & Gilboa, 2014; van Kesteren et al., 2012; Webb et al., 2016). Ghosh and Gilboa (2014) recently proposed specific features that define a schema: (1) an associative network structure that represents units of information and their interrelations, (2) are based on multiple episodic events, (3) lack specificity in unit details, and (4) have a degree of adaptability. Due to the way schema has been applied throughout research, it has a broad definition and has been used to refer to other types of memory representation (e.g., gist and concepts). Therefore, despite the definition provided by Ghosh and Gilboa (2014), it is necessary to clarify what the present thesis considers "schema".

Gist-based representations are also an abstract representation, but only of a single episode (Nadel et al., 2000; Winocur & Moscovitch, 2011). An example of the difference between gist and schema can be demonstrated based on how both would represent a "party". For a gist, the representation will be isolated to one instance, such as your 18th birthday party. Here, more specific details about the event itself (e.g., there were balloons around the house) will be available, but without the need for explicit details (e.g., the balloons were red, yellow, and blue). In contrast, a schema will be based on several instances of attending a "party"; this will lead someone to anticipate what will be present in a future instance (e.g., cake, friends, cards, gifts), but without being based on a unique episode. Therefore, there are qualitative differences between schema and gist. However, both may co-exist and be used based on task demands (Robin & Moscovitch, 2017).

Concepts are another type of mental representation that capture the shared properties between similar items and experiences. There is an ongoing debate within the literature about the distinguishing features between concepts and schemas (Ghosh & Gilboa, 2014; Preston et al., 2017). The present work does not set out to clarify this debate but understands it would be challenging to disentangle schema and concepts given they share common features and underlying processes. For instance, behavioural work has shown that information congruent with schematic knowledge enhances recall (e.g., Brewer & Treyens, 1981), with similar error-driven enhancement shown for concepts (e.g., Sakamoto & Love, 2004). Further, in the neuroimaging literature, the hippocampus and medial prefrontal cortex (mPFC) have been implicated as key neural areas that are involved in schema and concept formation and use (Davis et al., 2012; Mack et al., 2016; Zeithamova et al., 2016). Therefore, there is a great degree of overlap between schema and concepts.

The present thesis does not aim to address whether schema, gist and concepts are distinct. Instead, by presenting this information, it should be apparent how contentious an issue it is to define the term “schema”. For the present work, many schema features, as discussed by Ghosh and Gilboa (2014), are present (this is discussed in more detail in Defining Schema in the Precision Paradigm). However, research tied to concepts may also be relevant to the present work, with the global term “schema” used throughout the thesis to encapsulate both forms of representation.

1.2.2. Schema and Memory

1.2.2.1. The Advantages of Schema Presence

Schemas have been shown to benefit memory when information is either congruent (Aizpurua et al., 2009; Atienza et al., 2011; Bower et al., 1979; Brewer & Treyens, 1981; Mandler & Johnson, 1977; Nakamura et al., 1985; van Kesteren et al., 2010; Yamada & Itsukushima, 2013) or incongruent (Frank et al., 2018; Greve et al., 2019; Hunt & Worthen, 2006; Tulving & Kroll, 1995) with schematic information. Brewer and Treyens (1981) had participants recall items present in an office they were asked to wait in for 35 seconds. Items in the office were either congruent (i.e., items expected given the context), such as a desk, or incongruent (i.e., items that would be unusual given the context), such as a picnic basket. In the study, schema-congruent items were found to be better remembered than incongruent items. Such a congruency benefit has been shown across several paradigms, such as: item-colour pairings (Cycowicz et al., 2008), item-location pairs (Atienza et al., 2011; Tompary & Thompson-Schill, 2021), word lists (Packard et al., 2017), stories (Mandler & Johnson, 1977) and films (van Kesteren et al., 2010). Therefore, the congruency benefit is a robust phenomenon.

Nevertheless, despite memory performance for schema-congruent information being typically greater than for schema-incongruent information, research also suggests that schemas can boost memory for schema-incongruent information relative to unrelated information. Frank et al. (2018) had participants learn a set of events that produce a coherent representation (e.g., A-B, B-C, C-D, D-A). The first pairing (i.e., A-B) providing schema context, the second (i.e., B-C) included schema-congruent or -incongruent information, whilst the final element (i.e., D) was always schema-congruent with the overall event. For example, the A-B pairing could be Farm-Tractor, with C then being either congruent (e.g., Farmer) or incongruent (e.g., Lawyer). These two conditions were compared to a control condition where all item pairings were unrelated (e.g., Torch-Professor, Professor-Lego). Consistent with Brewer and Treyens (1981), memory performance was greater in the schema-congruent relative to -incongruent condition. However, they also saw (in some circumstances) greater memory performance in the schema-incongruent relative to the unrelated control condition. Consequently, schemas may benefit the encoding and retention of congruent and incongruent information under specific conditions.

The schema-linked interactions between medial prefrontal and medial temporal regions (SLIMMs) model (van Kesteren et al., 2012) proposes that the mPFC and hippocampus play important roles in schemas, affecting both memory encoding and retrieval. According to SLIMMs, when an event is congruent with present schematic representations, mPFC activation occurs to ensure the new event is rapidly integrated with existing neocortical schemas. The mPFC does this by inhibiting hippocampal processing for that event, thus suppressing the encoding of schema-irrelevant information (e.g., perceptual details). In contrast, when information is incongruent with schematic representations, the mPFC does not suppress hippocampal encoding. As such, the entire experience is encoded, including contextual details that may be irrelevant to the schema it is incongruent to. Allowing all details to be encoded allows for identifying what factors resulted in the prediction error, allowing for schema updating. Therefore, this model proposes that schema congruence enhances recall due to the new event being embedded with neocortical representations. In contrast, incongruent events are recalled with greater precision due to the prediction error caused, leading to better encoding of the material.

In their study, Greve et al. (2019) set out to test the proposals of the SLIMMs model. Here, participants learned the value of a set of objects via trial-and-error learning. Two objects would appear on screen, with particular objects (e.g., umbrella) having higher values than others (e.g., shoes). Across learning, the value of an item may have remained the same (congruent trials), changed on the final trial (incongruent) or had no fixed value and changed on every trial (unrelated). At test, participants were asked to identify whether an object display was old or new and which object had the higher value. Across four experiments, it was found that recognition was better for congruent and incongruent information compared to unrelated information; this supports the findings of Frank et al, (2018) above. However, these findings then extend the Frank et al. (2018) findings. Specifically, Greve et al. (2019) found an advantage to recognising first-encountered items in the congruent trials despite having no distinguishing characteristics at this point; this suggests the congruency benefit of these trials enhanced memory post-encoding. For the incongruent trials, where the final trial changed the object's value from previous learning, recognition was greater than in the unrelated information condition. This finding supports the predictions of the SLIMMs model (van Kesteren et al., 2012). Specifically, the presence of incongruent information may result in prediction error, thus enhancing encoding for those items and leading to better memory at test as items violate expectations.

An important consideration is the time it may take for a schema to develop before violations to patterns affect memory processing. Richter et al. (2019) had participants learn associations between locations and objects around a circle. Object-locations were associated with a particular area of the circle (e.g., clothing with the top left-quadrant). However, the schema changed for one group of items such that locations were rotated by 90° (e.g., clothing was now associated with the top right quadrant), though this happened at different learning points. For the consolidation group, participants did not encounter the new schema information until the subsequent day of training. In contrast, for the no-consolidation group, this schema change occurred on the same day. It was found that those in the consolidation group showed a greater ability to update their existing schema to fit with the new object-location pattern. In contrast, those in the no-consolidation group showed difficulty shifting from the old (e.g., top left quadrant) to the new (e.g., top right quadrant) pattern. This result suggests that schema updating requires the initial stabilisation of the newly learned information before schema-inconsistent information can be identified. Applying the SLIMMs model to the no-consolidation group, as the schema had not yet undergone consolidation, the presence of schema-inconsistent information did not violate expectations. However, once the schema was developed, it relied on the initially learned experiences where most attention would have been devoted. In contrast, when the schema had time to develop (i.e., the consolidation group), the schema could identify schema-inconsistent information and update appropriately via prediction error mechanisms.

1.2.2.2. The Disadvantages of Schema Presence

All the studies presented above suggest that schemas can benefit memory. However, there are disadvantages present when memories are affected by schema presence, leading to false recall or inaccuracies in the reported events (Aizpurua et al., 2009; Bartlett, 1932; Berens et al., 2020; Bower et al., 1979; Brewer & Treyens, 1981; Garcia-Bajos et al., 2009; Lampinen et al., 2001; Lew & Howe, 2017; Nakamura et al., 1985; Yamada & Itsukushima, 2013). In the classic example described earlier, Bartlett (1932) found that participants often changed story elements to fit their understanding of the world. In their study, Brewer and Treyens (1981) found that some items were reported to be present (e.g., stationary) that were not in the room simply due to schema-expectancy. A classic paradigm to illustrate false memories resulting from schema is the Deese-Roediger-McDermott paradigm (Deese, 1959; Roediger & McDermott, 1995). In this task, participants are asked to learn a list of words that share a semantic relationship (e.g., sleep, tired, bedroom). Subsequently, when asked to remember the items learned, participants will often remember items that were semantically associated (e.g., pillow) but not present in the lists. Evidently, the interrelation among stimuli may have influenced schematic processes, leading to false memories for words that were not studied.

These behaviours are also observed in other types of task, such as eyewitness reporting (Loftus et al., 1978), and word- or object-location associations (Berens et al., 2020; Lew & Howe, 2017; Richter et al., 2019). In their study, Brady et al. (2018) investigated how schema may bias the reporting of events. Here, participants learned to associate objects with a particular colour. Objects came from four different categories, with each category having a specific colour. For instance, backpacks were associated with red(ish) colours. The colours selected were based on a von Mises distribution (i.e., circular Gaussian), meaning the primary colour (e.g., red) was more likely than colours further away (e.g., purple). They found that participants were more likely to report the colours of objects as closer to the mean of their given category; this shows how the presence of a pattern may bias memory for events.

The presence of a schematic pattern has also been shown to affect forgetting differentially (Berens et al., 2020). In their study, Berens et al. (2020) had participants learn word-location associations around a circle. Words belonged to two semantic categories: human-made (e.g., chair, table) and natural (e.g., apple, giraffe). Unbeknownst to participants, one group of words had locations clustered in one area of the circle (the clustered condition). The other had no pattern underlying the locations associated with the words (the non-clustered condition). Using this method, Berens et al. (2020) were able to investigate memory accessibility (i.e., the proportion of items remembered) and precision (i.e., the degree of error from the location presented at study to the one selected at test). Interestingly, the presence of a pattern (or schema) differentially influenced these metrics of memory performance. While memory accessibility was greater in the clustered than the non-clustered condition, the opposite was true for memory precision, whereby the non-clustered items were more precisely remembered than the clustered. These differences were found to remain over a 96-hour delay period. This study demonstrates both the positive and negative influence that schematic information may have on memory. While more information may be accessible, the items themselves are likely to be reported with a greater degree of error than items with no schematic information.

1.2.3. Summary

Schemas represent the abstracted regularities across experiences and share features with other forms of memory representation (e.g., gist and concept). Previous work has shown that schemas can positively affect memory, with better memory when items are congruent or incongruent with schema than schema-irrelevant material. However, there are also disadvantages to schema presence. Specifically, schemas can distort memory, leading to increased false alarms and reduced memory accuracy.

1.3. Memory-based Generalisation

Though the review's focus has been on remembering past events, memories are formed to guide future behaviour. Memory-based generalisation is central to the present thesis and has been studied using several methodologies: inductive reasoning (Rips, 1975), associative inference (Carpenter & Schacter, 2017; Kumaran, 2013; Preston et al., 2004; Zeithamova et al., 2012), affix generalisation (Tamminen et al., 2012), word-picture matching (Mirković & Gaskell, 2016), weather prediction (Kumaran et al., 2009), face-location paradigms (Sweegers & Talamini, 2014), and object-location tasks (Tompary et al., 2020).

1.3.1. Studies of Memory-based Generalisation

Using relational reasoning paradigms, Rips (1975) demonstrated how participants rely on existing knowledge to generalise to a novel instance. In their study, Rips (1975) asked participants to identify the likelihood that animals on an isolated island would develop a pathogen. It was found that participants grouped the animals based on their perceived similarity (e.g., eagle and hawk). When this occurred, they were more likely to give birds a greater likelihood of possessing the pathogen, the greater the perceived similarity to the first carrier species. For example, if hawk were disclosed as the original carriers of the pathogen, eagles would have a higher likelihood of also being a carrier than geese or ducks. Here, individuals rely on both the presented information and their semantic knowledge regarding the relatedness of each animal to make estimates of pathogen likelihood.

In their study, Sweegers and Talamini (2014) required participants to learn face-location associations, with certain facial feature combinations predictive of locations. During study, participants observed a face moving to one of six possible locations organised hexagonally around the centre of the screen. They were then asked to retrieve these locations after each block of trials, acting as a form of retrieval practice. For the study, there were three groups: immediate, no-nap and nap, with the latter two groups having a 4-hour period between study and test. During test, participants were presented with both old and novel faces. It was found that participants were capable of generalising above chance levels, with the no-nap group showing significantly better performance than the immediate group, with no differences between the nap and no-nap groups nor the nap and immediate. During a one-month follow-up, the rates of forgetting were found to significantly differ, with most forgetting occurring for faces that did not have a rule for their location based on facial features. However, despite evidence of forgetting, the ability to generalise did not significantly change. These results reiterate the earlier proposals about schema and forgetting – specifically, maintaining more accessible information when a schema is present than when it is not (Berens et al., 2020).

More recently, Tompary et al. (2020) investigated memory-based generalisation to assess schema use over time. This study followed a similar procedure to what was described in Berens et al. (2020), but with some notable differences. First, in the Tompary et al. (2020) study, participants were presented with object-location pairings and not word-location pairings. Second, there were two clusters present in the Tompary et al. (2020) study, with the means of these clusters being 180° apart. In contrast, the Berens et al. (2020) study only implemented a cluster in one of the conditions. The other condition (the non-clustered) did not have an underlying pattern associated with the word locations. Finally, in Tompary et al. (2020), participants were made to generalise their learning to novel instances; this occurred either 24-hrs or 1-week after encoding. No generalisation trials were explicitly used in the Berens et al. (2020) study. Using this paradigm, Tompary et al. (2020) found that participants were biased towards the mean representation of events, similar to the Brady et al. (2018) result. Additionally, participants could generalise the patterns presented at study during test, both at 24-hrs and 1-week later. However, adherence to the pattern presented declined over time. In their study, Tompary et al. (2020) noted that schema use increased with time (as evidenced by an increased tendency to report items as closer to their mean). However, these schematic representations did decay as memories for the individual events themselves decayed. More specifically, using the generalisation trials, it was found that participants were increasingly less likely to remap the pattern presented at study after a week compared to 24-hours post-encoding. Further, memory for older items tended to be better when items were more schema consistent (i.e., closer to the centre of the cluster) than inconsistent (i.e., further from the centre of the cluster). These contrasting findings suggest that over time there is a loss of information for the underlying distribution for a given group, despite less error occurring in the reporting of items that were more schema-congruent than incongruent. These results suggest participants can generalise after a period of consolidation (24-hrs between study and test), with this ability declining over time.

However, the discrepancy in outcomes for the Sweegers and Talamini (2014) and Tompary et al. (2020) studies are worth noting. Specifically, whilst Sweegers and Talamini (2014) found that generalisation remained stable over a month's delay, Tompary et al. (2020) found that over a 1-week delay, generalisation declined. One explanation for this discrepancy may relate to the sensitivity of the paradigms. Whilst Sweegers and Talamini (2014) used a binary measure of "correct" and "incorrect" generalisation, the paradigm used by Tompary et al. (2020) allowed for a more detailed examination of the patterns used by participants based on the locations selected around the entirety of the circle. As a result, the greater specificity provided by the Tompary et al. (2020) method may have meant that changes in generalisation were easier to identify. An alternative explanation for these discrepancies may relate to the approach used by participants in order to generalise. More specifically, whether an encoding- or retrieval-based form of generalisation was used; these models are discussed in more detail below.

1.3.2. Theories of Generalisation: Encoding vs Retrieval-Based Processes

Within the literature, there are two predominant theories for explaining memory-based generalisation: encoding-based (also referred to as prototype models, e.g., McClelland et al., 1995; Posner & Keele, 1968; Smith & Minda, 2000) and retrieval-based (also referred to as exemplar models, e.g., Kumaran & McClelland, 2012; Nosofsky, 1988; Shohamy & Wagner, 2008). Most encoding-based models propose that as experiences are encoded, the overlapping patterns of information form a schematic representation of the central tendencies (e.g., "average") of these items (Rosch, 1973). In contrast, retrieval-based models propose that the individual experiences (or exemplars) are used "on the fly" to make an inference (Hintzman, 1986; Medin & Schaffer, 1978; Nosofsky, 1986). For example, recalling three related experiences and using them to generalise to a novel instance. The critical difference between encoding- and retrieval-based models is how memories are used. For encoding-based models, schematic representations are formed during or after encoding as a function of systems consolidation, meaning there is less reliance on individual experiences. In contrast, retrieval-based models argue that individual experiences are sampled from memory and used to make an inference.

Though evidence supports both encoding-based and retrieval-based models (see Nosofsky, 1988 and Smith & Minda, 2000 for meta-analytic reviews), they are not necessarily mutually exclusive (Medin et al., 1984). Instead, arguments have been made that different generalisation mechanisms may be used under different task conditions. For instance, the more variability there is among items, the more need for retrieval-based mechanisms to discriminate among items within a category (Konkle et al., 2010; Winograd, 1981). However, the more overlap across events, the less need for individual episodes, meaning encoding-based generalisation may be sufficient (Hampton, 1979; Rosch & Mervis, 1975).

As a result of this shift in perspective, more recent neurocognitive models have proposed how both encoding- and retrieval-based mechanisms may develop. For example, the recurrency and episodic memory results in generalisation (REMERGE; Kumaran & McClelland, 2012) model, which builds on the CLS model (discussed earlier), proposes that the hippocampus initially supports generalisation before systems consolidation. Specifically, the hippocampus will use big-loop recurrence to use its output as input. This "new" input will activate related experiences, which can be recombined into a generalised representation of related events. These generalised representations can then be used to generalise to novel instances. However, throughout consolidation, there may be a shift from retrieval-based to encoding-based generalisation, as argued by the CLS model. This hybrid model could accommodate findings that show schema effects immediately, or soon after, learning (e.g., Antony et al., 2021; Sweegers & Talamini, 2014; Tompary & Thompson-Schill, 2021) and the slow maturation of neocortical schematic representations, which can take days, weeks, months, or years (McClelland et al., 1995; Zola-Morgan & Squire, 1990). Based on the evidence reviewed so far, it is clear that both encoding- and retrieval-based mechanisms may operate at different time scales and under different conditions.

Going back to the earlier Tompary et al. (2020) and Sweegers and Talamini (2014) discrepancy, different generalisation mechanisms may have been at play. Specifically, during the Tompary et al. (2020) investigation, retrieval-based generalisation may have been used; this is in line with their finding that "schema" use (operationalised as greater reliance on the mean of each cluster) declined as the memory for items themselves declined. Therefore, generalisation appears to have resulted from reliance on the individual experiences instead of a schematic representation. This is particularly evident when considering that schema are meant to be long-lasting mental representations independent of the memories they are based upon (Ghosh & Gilboa, 2014; van Kesteren et al., 2012). In contrast, for the Sweegers and Talamini (2014) study, the relative stability of generalisation following a one-month delay despite declines in memory suggests the presence of a schematic (encoding-based) representation. The development of a schematic representation may have occurred as a result of the retrieval practice trials during learning that may have encouraged online consolidation (see Antony et al., 2017, for a discussion on consolidation via retrieval practice).

1.3.3. Summary

Throughout this section, it has been shown how memories can be used to make inferences about novel stimuli. Many paradigms have been used to test this, often requiring a binary output (e.g., associative inference and face-location task). However, there has been a move towards measures that allow for greater insight into the pattern extracted by participants, resulting in a continuous output of locations around a circle (i.e., Tompary et al., 2020). Both encoding and retrieval-based models were discussed, showing that either mechanism may influence behaviour based on task demands or the influence of time.

1.4. Neural Mechanisms

The review now turns to the neural mechanisms of memory-based generalisation. In the final experimental chapter of the thesis (Chapter 4), these neural mechanisms were explored in a pilot investigation. As mentioned above, many studies and theories have highlighted the importance of the ventral mPFC (vmPFC) and hippocampus in memory-based generalisation, particularly in their interaction (Andrews-Hanna et al., 2010; Kumaran &

McClelland, 2012; McClelland et al., 1995; Schlichting & Preston, 2015; van Kesteren et al., 2012; Zeithamova et al., 2008). For example, the SLIMMs model proposes that schema-congruent information becomes rapidly consolidated into neocortical structures through mPFC suppression of hippocampal encoding. In contrast, schema incongruent events will result in no such suppression of hippocampal activation, allowing for encoding of the unique experience to take place.

The hippocampus is regarded as serving a role in the rapid learning of events (McClelland et al., 1995), with recent evidence also demonstrating its role in generalisation (Bowman & Zeithamova, 2018; Preston et al., 2004; Shohamy & Wagner, 2008; Zeithamova et al., 2012). In contrast, the vmPFC has been shown to be involved in several schema related processes such as schema updating (Richards et al., 2014; Zeithamova et al., 2012), enhancing memory for schema-congruent information (Tse et al., 2007; van Kesteren et al., 2010) and generalisation (DeVito et al., 2010; Kumaran et al., 2009; Schlichting et al., 2015).

1.4.1. The Hippocampus and vmPFC

Historically, studies have suggested that the hippocampal and vmPFC activity can serve as ways of dissociating the form of generalisation process utilised by participants (Bowman & Zeithamova, 2018). Specifically, hippocampal activation was associated with retrieval-based generalisation, whilst the vmPFC was associated with more encoding-based processes. This dissociation is partly due to the hippocampus' role in episodic memory, meaning the region can access recently encoded experiences and generalise based on the individual episodes; this is central to the REMERGE model (Kumaran & McClelland, 2012). In contrast, schema are developed and maintained via the vmPFC, with inferences about novel situations being possible via encoding-based processes using these schematic representations (Zeithamova et al., 2012).

However, recent proposals have been made that the hippocampus represents events differently across the long axis, with posterior regions more involved in memory and anterior regions more involved in generalisation (Frank et al., 2019; Kumaran & McClelland, 2012; Poppenk et al., 2013). One area of support for this proposal comes from animal work. The receptive fields of place cells in the rodent hippocampus have been shown to change across the hippocampus, with smaller place fields in posterior and larger place fields in anterior portions (Kjelstrup et al., 2008). Thus, more anterior portions of the hippocampus may represent information on a larger spatial scale, potentially allowing for generalisation. In the human literature, fMRI studies have corroborated this notion by showing that the hippocampus can be differentially active for memory compared to generalisation trials (Bowman & Zeithamova, 2018; Collin et al., 2015; Schlichting et al., 2015).

Schlichting et al. (2015) had participants learn associations among objects (e.g., A-B, B-C). Before and after learning, participants were exposed to each object individually; this took place within an fMRI setting. The reason for collecting neural data on the response to singular objects was to allow for a representational similarity analysis (RSA) to be conducted. Specifically, RSA was used to test specific hypotheses about memory separation and integration. For instance, to identify differences in hippocampal activation based on the integration of conditions (i.e., A-B-C) and separation of learning (A-B, B-C). Post-scanning, participants took part in an associative inference task where they would have to infer the relationship across unexposed images (e.g., A-C) along with remembering the original pairs (e.g., A-B, B-C). It was found that the posterior region of the hippocampus was associated with learned pairs (e.g., A-B, B-C), whilst unlearned novel pairs that shared an associative relationship (e.g., A-C) were shown to activate anterior regions of the hippocampus. This was demonstrated as posterior hippocampal regions showed less similarity for A and C items post-learning. In contrast, anterior regions of the hippocampus showed greater pairing post-learning.

The notion of hippocampal representations that allow for generalisation was briefly touched upon when discussing the REMERGE (Kumaran & McClelland, 2012) model. Using the associative inference paradigm as an example, REMERGE proposes that the hippocampus will code for each unique experience (e.g., A-B, B-C, C-D) via pattern separation mechanisms. Subsequently, as the hippocampus uses recurrence to transfer information to the neocortex, some information will feedback into the hippocampus proper via the entorhinal cortex. This feedback loop is termed "big-loop recurrence". The feedback received will act as a new input into the hippocampus and activate related experiences. For instance, if the B-C association acted as the new input, it would activate other conjunctive experiences with both B and C present (i.e., A-B, B-C, C-D). This subsequent re-activation will allow the hippocampus to identify higher-order relationships across events that may not have been directly experienced (e.g., A-C, B-D). The hippocampus can then store these generalised representations that represent higher-order relationships across events. During generalisation, the hippocampus can then use these generalised representations to infer the relationship across events. At times, new input from a novel experience will also activate associated information, thus leading to an "on the fly" inference to be made at the time of generalisation if a generalised representation were not present.

In their study, Bowman and Zeithamova (2018) trained participants to distinguish between two categories of stimuli based on specific features present in exemplars (e.g., head direction, tail appearance). At test, participants had to classify both old and novel stimuli into one of the two categories. Models were then fit to participant data to identify whether an encoding- or retrieval-based approach was used when generalising to novel instances. It was found that the encoding-based model provided the best fit for the behaviour of participants in around 72.4% of cases, compared to 10.3% for retrieval-based processes. The models could not distinguish the remaining 17.2% of participants. Interestingly, when comparing the model fit for each participant to the neural activation found, the encoding-based model fits correlated with vmPFC and anterior hippocampus activation, suggesting the regions contributed to encoding-based generalisation. However, no regions were found to track retrieval-based mechanisms unless the statistical threshold was reduced. When this occurred, retrieval-based processes were associated with bilateral occipital cortex, precuneus, and inferior partial cortex. These findings may also be in line with the expectations of the REMERGE model, potentially highlighting the generation of generalised representations within the anterior hippocampus and schema-related processes in the vmPFC.

Further support for the interaction between the hippocampus and vmPFC during generalisation was found in Frank et al. (2019). In their study, Frank et al. (2019) found connectivity differences between the posterior and anterior regions of the hippocampus. Specifically, the posterior region showed greater connectivity to regions previously implicated in memory specificity (e.g., angular gyrus and inferior frontal gyrus), whilst anterior regions showed greater connectivity with the vmPFC. The interaction between the hippocampus and vmPFC was also predictive of individual differences in generalisation ability. This result further demonstrates the importance of both the vmPFC and hippocampus in memory-based generalisation. It also provides further weight to the notion that sub-regions of the hippocampus may serve dissociable processes.

However, it should be noted that not all studies find model-based differences favouring encoding-based models (see Mack et al., 2013). In their study, Mack et al. (2013) applied multivoxel pattern analysis to fMRI data collected during a categorisation task; this was done to determine whether an encoding or retrieval-based model provided the best fit for the data. In this study, participants were trained to classify objects as belonging to one of two categories based on binary differences (e.g., when red, it is category A, when green, it is category B). During scanning, participants classified both old and novel objects. Using multivoxel pattern analysis and support vector regression, it was found that a retrieval-based model provided the best fit for the neural data, with 65% of participants fitting the retrieval-based model, 5% fitting the encoding-based model, and the remaining sample not being dissociable based on brain activation (30%). This finding seems at odds with the Bowman and Zeithamova (2018) study. However, Mack et al. (2013) found that the regions associated with retrieval-based processes differed from the encoding-based regions found in Bowman and Zeithamova (2018). Specifically, Mack et al. (2013) found that the lateral occipital, inferior parietal cortex, inferior frontal gyrus, and insular cortex were key in retrieval-based generalisation. This result does indicate there may be neural dissociations between encoding- and retrieval-based generalisation. Along with this, the findings give merit to earlier behavioural discussions that different tasks may use either encoding or retrieval-based processes. Though both tasks required categorisation of stimuli, the stimuli used were more complex in the Bowman and Zeithamova (2018) study and may have driven more encoding-based generalisation processes to be required. In contrast, the simpler stimuli of Mack et al. (2013) may have required the use of retrieval-based processes where examples of the stimuli can be retrieved and used to infer a category.

1.4.2. Beyond the hippocampus and vmPFC

Other regions outside the hippocampus and vmPFC have also been implicated in both memory and generalisation. For instance, the middle temporal gyrus has been found to support generalisation (Bowman & Zeithamova, 2018; Davis & Poldrack, 2014; Dennis et al., 2008; Turney & Dennis, 2017). Previously, work had also implicated the striatum as a critical region for non-declarative memory-based generalisation (Knowlton et al., 1996). Studies have also shown how different memory processes may be dissociable on a neural level. In their study, Richter et al. (2016) found that hippocampal activation predicted memory accessibility, angular gyrus activation predicted memory precision and precuneus activation predicted vividness. Consequently, studies have previously demonstrated that regions outside of the vmPFC and hippocampus may play important roles in both memory and generalisation.

1.4.3. Summary

The predominant focus of the literature on memory-based generalisation has been on the vmPFC and hippocampus. Both have been shown repeatedly to be integral regions to this cognitive domain, both separately and in their interaction. More recently, there has been a shift to looking at the division of labour along the long axis of the hippocampus, finding that the anterior portion may be more involved in generalisation whilst the posterior is involved in memory. The coupling between these portions of the hippocampus and other regions also supports this notion. Other areas have also been implicated in memory and generalisation but have been given less attention (e.g., middle temporal gyrus).

1.5. Thesis Overview

The present review has demonstrated that the presence of a schema affects the encoding and retention of related events. Typically, schemas improve memory performance for congruent and incongruent information, relative to schema-irrelevant information (Frank et al., 2018; Greve et al., 2019). However, less is known about the effects of a schema on information that is irrelevant. If we experience multiple related events that are intermixed with unrelated events, does the extraction of a schema for the related events affect performance for these unrelated, irrelevant, events? Current theories do not make clear predictions about schema-irrelevant information (Henson & Gagnepain, 2010; McClelland et al., 1995; van Kesteren et al., 2012), though most would assume that such events should be unaffected by the presence of a schema. If information is unrelated to a schema, then the schema should not modulate its encoding or retention.

Schemas can relate to the locations of items in the real world. When entering someone's home, semantically related items are typically grouped spatially. If you know where the soap and toilet paper is located, you can use this information to predict where a towel will be located. However, this schematic information will be of little use when predicting the location of items from an unrelated category. For example, house plants can be placed anywhere in a home, and as such, the presence of a "bathroom" schema should be irrelevant to where a house plant is located or where you predict one might be.

The present thesis had three primary aims: (1) explore memory-based generalisation with a method that allowed closer inspection of the patterns extracted and used by participants, (2) investigate whether schematic information influences behaviour for both relevant and -irrelevant information and (3) assess the neural mechanisms of memory-based generalisation through fMRI. Therefore, this next section provides specific details regarding the remainder of the thesis. First, there is a discussion on the precision paradigm used throughout the thesis to understand memory-based generalisation and why this was chosen. Subsequently, individual overviews of each Chapter are given to provide further insight into the thesis overall.

1.5.1. The Precision Paradigm

Precision (or continuous) measures of memory provide a non-binary output that allows us to look at patterns of responses across trials. This paradigm has been mentioned in passing throughout the present review and has been used extensively to study working (Bays et al., 2009; Luck & Vogel, 2013; Peich et al., 2013; Sun et al., 2017; Zhang & Luck, 2008) and long-term (Berens et al., 2020; Harlow & Donaldson, 2013; Korkki et al., 2020; Nilakantan et al., 2018; Richter et al., 2019; Tompary et al., 2020) memory. Precision memory experiments associate a stimulus (e.g., a word or object) with a continuous property (e.g., colour or location around a circle). At test, participants are required to retrieve the associated property of the stimulus. In the case of a location around a circle, performance is measured as the degree of error (real location vs. retrieved location). With a continuous measure like this, we can assess the distribution of retrieved locations across trials. For example, we can compare the distribution of memory trials for a set of stimuli whose locations were dictated by an underlying pattern to see if the retrieved distribution matches the encoded distribution.

The present thesis adapted the method used by Berens et al. (2020). As a recap, Berens et al. (2020) examined forgetting using the precision paradigm. Two semantic word groups were used: human-made and natural. One set of words had locations that were more likely to appear in one area of the circle than elsewhere (the clustered condition). The other had locations that were equally likely anywhere around the circle (the non-clustered condition). One adaption was made for the present work to investigate memory and generalisation using this paradigm, specifically the inclusion of novel words at test. These words came from the same semantic groupings but were not associated with a location around the circle. Instead, participants are required to determine their location based on previous experience. Therefore, the present thesis could test whether participants could recreate the pattern of locations observed at study for novel items. Using an approach that assesses the pattern of responses across trials also allows us to identify biases more clearly in behaviour; this is particularly relevant for addressing the second aim of the present thesis (i.e., identify whether schematic information may bias behaviour for schema-relevant and -irrelevant information).

Much of the previous literature on memory-based generalisation has used tasks requiring binary decisions (e.g., Knowlton et al., 1996; Kumaran et al., 2009; Preston et al., 2004; Sweegers & Talamini, 2014; c.f. Tompary et al., 2020) that address whether memory-based generalisation is possible. However, the precision paradigm provides insight into the patterns extracted and used during the recall of events and generalisation to novel instances. Therefore, it provides a more sensitive measure into memory-based generalisation than is typically offered by these binary measures. For instance, in Tompary et al. (2020) it was possible to track the development and use of a schema through the precision paradigm as demonstrated by decreases in the use of the underlying pattern associated with a particular cluster. In contrast, for studies such as Sweegers and Talamini (2014), it can be difficult to discern subtle changes in behaviour as items are either correctly or incorrectly located without examination of the precise patterns being used by participants.

An additional benefit of this paradigm is that it does offer the opportunity to construct a schema (this is discussed in more detail below). Unlike other methods which rely on pre-existing schema (e.g., semantic association between objects), the precision paradigm can construct a schema by associating particular spatial locations with a category of words or objects. Though there will be use of pre-existing knowledge (e.g., semantic association among stimuli) they will not possess information related to a spatial distribution. As such, associating both word and locations together provides a useful way of examining schema development and use.

The present work will also extend previous work conducted by Tompary et al. (2020) due to the inclusion of a control (non-clustered) condition. In their study, Tompary et al. (2020) had two clusters on opposite sides of the circle. Though this was beneficial for examining schema development and use over time, it did not allow for analysis of how a pattern (or schema) in one condition influenced both itself and other conditions. Therefore, the presence of a non-clustered condition within the present paradigm would allow for the extension of Tompary et al. (2020)'s findings and address how schema presence influences behaviour for relevant and irrelevant items.

1.5.1.1. Defining Schema in the Precision Paradigm

Critical to the present thesis is the use of an underlying pattern across a set of word-location associations to provide insight into schematic processing. As discussed earlier, Ghosh and Gilboa (2014) recently proposed specific features for a schema: (1) an associative network structure, (2) based on multiple episodes, (3) lack specificity, and (4) have a degree of adaptability. Concerning the present paradigm, a participant may rely on a schema that maps the associations between words and locations (related semantically and by location), thus conforming to the first criterion of an associative network structure. Further, participants are encoding multiple events, conforming to the second criteria. If a pattern is extracted (e.g., the average word-location association for a given semantic category), this conforms to the third criteria. Finally, though "adaptability" is not assessed per se (i.e., the extent to which existing schema can be flexibly updated), the work does assess behaviour shortly after encoding. Therefore, if behaviour is consistent with schema processing, schematic representations must have been developed rapidly. Consequently, the work presented within the thesis does conform to the stringent criteria outlined by Ghosh and Gilboa (2014) and readily fits with less stringent definitions of schemas (see Preston & Eichenbaum, 2013 and van Kesteren et al., 2012).

1.5.2. Thesis Chapters

1.5.2.1. Chapter 2

Previous work has often focused on how schematic information influences recall for congruent and incongruent information. However, little is known about how these schematic representations influence behaviour for irrelevant information. This is particularly evident in many theories that discuss schema development (e.g., SLIMMs), where no explicit assumptions are made regarding the effects of schema on schema-irrelevant information. Notably, this may be due to the implicit assumption that information irrelevant to the schematic representation should be unaffected by its presence. However, Chapter 2 sought to test this assumption by explicitly examining the pattern extracted in both the schema-relevant (clustered) and irrelevant (non-clustered) conditions through the scope of generalisation trials. Returning to the bathroom schema example, the towel could be in the bathroom with other bathroom-related items (schema-congruent) or in the living room (schema-incongruent). Whereas the location of the towel could either be schema-congruent or -incongruent, the location of a specific house plant is schema-irrelevant, as it should not be included as part of the “bathroom” schema. For the present paradigm, you have both schema-relevant (clustered) and schema-irrelevant (non-clustered) items. Therefore, it is possible to assess how schema presence affects behaviour towards relevant- and irrelevant information.

Four experiments were conducted. In Experiment 1, an assessment of whether generalisation could occur immediately post-encoding took place. Experiment 2 extended the findings from Experiment 1 by introducing a delay between study and test to assess whether there were changes in generalisation behaviour as a function of time. Subsequently, Experiments 3 and 4 (which were conducted online) aimed to replicate the effects observed in the first two experiments to assess the robustness of the effects presented. All experiments were pre-registered before data collection.

1.5.2.2. Chapter 3

Chapter 3 aimed to better understand the pattern of behavioural effects observed in Chapter 2. Computational modelling was used to address whether: (1) an encoding- or retrieval-based model predicted the pattern of behaviour observed in Chapter 2, (2) modulating retrieval probabilities produces the patterns of behaviour from Chapter 2, and (3) alternative mechanisms (e.g., interference) could produce the same pattern of effects found in Chapter 2. These models provided the opportunity to explore the fit of encoding and retrieval-based models to the data obtained in Chapter 2 and explore alternative interpretations for the effects present in Chapter 2 and other related studies (e.g., Berens et al., 2020; Tompary et al., 2020).

1.5.2.3. Chapter 4

The final experimental chapter (Chapter 4) initially aimed to assess the neural correlates of memory-based generalisation via fMRI. However, due to the COVID-19 pandemic, data collection was halted and left incomplete. As such, the chapter now provides an analysis of two pilot investigations. The first was a behavioural pilot that assessed how the inclusion of a semantic categorisation task (SCT), aimed at reducing item-novelty effects present during generalisation trials, impacted the behaviour observed in the precision paradigm. The second was the preliminary analysis of the fMRI data to identify possible design and analysis changes and open questions that could be explored in an independent investigation.

1.6. Concluding Remarks

Schema have been shown to benefit memory when items are congruent and incongruent with expectations. However, an underlying assumption of many theories is that schema-irrelevant information is unaffected by schema presence. However, this notion has not been formally tested. Using generalisation trials, it is possible to ascertain how schema affect behaviour for information relevant and irrelevant to themselves. Specifically, these trials are not associated with an individual experience. Therefore, they provide an opportunity to examine schema-influence. However, many existing methods of testing memory-based generalisation rely on binary outcomes, meaning they are limited in what information can be provided about the patterns extracted and used by participants. Implementing the precision paradigm provides a novel opportunity to examine how the presence of a pattern (or schema) influences behaviour for both relevant and irrelevant information and the ability to examine the precise patterns used by participants over time. Therefore, the present thesis sought to assess how schematic information influenced behaviour, alternative explanations for the behaviour observed and gather preliminary evidence of the neural activations associated with memory-based generalisation.

Chapter 2:
Influence of schematic information on generalisation behaviour for both schema-relevant and -irrelevant information

All experimental pre-registrations, materials, data, and analyses are available on the Open Science Framework: https://osf.io/bxru4/.

The experiments presented within this chapter have been previously published as a preprint: Cockcroft, J. P., Berens, S., Gaskell, M., & Horner, A. J. (2021, August 24). *Schematic information influences memory and generalisation behaviour for schema-relevant and -irrelevant information*. https://doi.org/10.31234/osf.io/nzurq. Minor edits were made to ensure greater continuity between thesis chapters. However, the content is predominately the same as in the listed publication.

2.1. Abstract

It is typically assumed schema do not influence behaviour for information irrelevant to themselves. However, this has yet to be formally tested. Here, we assessed generalisation behaviour for information related to an underlying pattern, where a schema could be extracted (schema-relevant) and information that was unrelated and therefore irrelevant to the extracted schema (schema-irrelevant). To investigate this, the precision paradigm was used where participants learnt associations between words and locations around a circle. Words belonged to two semantic categories: human-made and natural. For one category, word-locations were clustered around one point on the circle (clustered condition), while the other category had word-locations randomly distributed (non-clustered condition). At test, participants were presented with old (memory) and new (generalisation) words, requiring them to identify a remembered location or make a best guess. The presence of the clustered pattern modulated memory and generalisation. In the clustered condition, participants placed old and new words in locations consistent with the underlying pattern. For non-clustered novel items, participants were less likely to place these items in locations consistent with the clustered condition. Therefore, we provide evidence that the presence of schematic information modulates memory and generalisation behaviour. In the case of schema-irrelevant information, the schema modulated generalisation behaviour. Our results highlight the need to carefully construct appropriate schema-irrelevant control conditions such that the presence of a schema does not modulate behaviour in these conditions.

Keywords*: schema, memory, generalisation*

2.2. Introduction

Schemas are mental representations that allow us to generalise across experiences, altering our memory of the past, perception of the present, and future predictions. Schemas are thought to be formed when we experience multiple related events that have a common structure (Anderson, 1984; Bartlett, 1932; Head & Holmes, 1911; Piaget, 1926; Posner & Keele, 1968). In this way, schemas may capture the general structure of events that have occurred, abstracting away from the specific content of individual events.

Schemas are thought to be critical to our ability to generalise to novel but related events. Sweegers and Talamini (2014) examined how the presence of an association between specific facial characteristics (e.g., wearing a hat, face shape) and a location in hexagonal space could be learned and used to make inferences for novel faces. Along with benefiting later recall of old items, the presence of face-location associations could also be used to make novel inferences about the location for unseen face stimuli. This was observed immediately after studying the material. In another domain, Mirković and Gaskell (2016) had participants learn new vocabulary using a word-picture matching task. When tested on their ability to generalise suffixes, participants showed they had extracted the suffix rules and were able to use these rules to generalise to novel word-picture pairs. Across these studies, it has been shown that schematic representations based on relational information can be used to make generalisations about novel stimuli. In this way, schemas do not simply function to benefit memory encoding and recall but also help guide our behaviour for future instances.

Models of schema processing can be broadly divided into: encoding-based or retrieval-based models. These groupings predominantly come from research into category and discrimination learning (Kumaran & McClelland, 2012; Mack et al., 2018; Murphy, 2016) but are relevant to schema processing. These are discussed in detail in Chapter 1. In brief, encoding-based models propose that at the point of encoding, extraction of regularities across events form schema that represent the central tendencies (e.g., average) across events (Rosch, 1973). In contrast, retrieval-based models propose that individual experiences are sampled at the point of generalising to guide novel inferences.

Precision memory measures have been used extensively to study working memory and long-term memory (see Chapter 1 for more details). Here, participants associate a stimulus (e.g., word or object) with a continuous property (e.g., colour or location around a circle). Participants then need to retrieve these locations, and their degree of error (i.e., difference from presented to selected location) can be used to assess memory accuracy. Precision measures have been used to assess schema processing. The idea here is that an underlying pattern can dictate the associated properties of a set of stimuli. For example, when learning word-location associations, the locations can conform to a von Mises distribution (circular Gaussian), such that they are clustered in a specific area of the circle; this was the approach taken in Berens et al. (2020). Here, participants were required to learn word-location associations around a circle. Words came from one of two semantic categories (i.e., human-made and natural), with one category having locations clustered (i.e., locations were more likely to appear in one area of the circle) while the other was non-clustered (i.e., no relationship between word meanings and locations). Using this paradigm, measures of memory accessibility (i.e., proportion of word-locations retrieved) and precision (i.e., degree of location accuracy given successful word-location retrieval) were assessed. They found that the presence of a pattern differentially influenced memory accessibility and precision. Specifically, accessibility was higher, but precision was lower, in the clustered relative to non-clustered condition. Consequently, schematic information affects distinct memory components differently – benefiting overall accessibility at the cost of precision.

Using both previously presented and novel (semantically related) stimuli, Tompary et al. (2020) investigated how these underlying patterns modulate both memory (old stimuli) and generalisation (novel stimuli) behaviour. Participants learned to associate objects with locations around a circle. The locations of images were drawn from two cosine distributions around the circle, with the means of these distributions being on opposite sides (separated by 180°). They found that schema use, relative to the use of episodic memory, increased with time, but interestingly, schema memory also showed evidence of decay. This is in line with evidence elsewhere showing that schema benefits on memory performance can decrease with time (Antony et al., 2021; Berens et al., 2020). Precision measures have therefore been used to assess memory and generalisation behaviour in the presence of a schema. However, they have not been used to assess behaviour for schema-irrelevant information.

Many studies that investigate schema processing consider its effects on schema-congruent and incongruent information, ignoring any potential influence on schema-irrelevant information. Irrelevant here relates to information that is not in the same semantic category as the schematic items. In our earlier example (see Chapter 1), the presence of a "bathroom" schema should have little impact on where you predict the house plant will be located (unless a mutual exclusivity rule is present; see General Discussion, below). Whereas a schema-incongruent item (e.g., a towel in the living room) conflicts with an existing schema and therefore can change or update the schema, a schema-irrelevant item (e.g., a house plant) is neither congruent nor incongruent with the schema. Therefore, our ability to remember where a house plant is located or predict where a house plant would be located should be unaffected by the presence or absence of schematic information related to bathroom items. In the case of the Berens et al., (2020) study, one semantic category (e.g., human-made – the experimental equivalent to bathroom items in our example) was associated with an underlying pattern (the clustered condition), whereas the other semantic category (i.e., natural – the experimental equivalent to house plants in our example) was not (the non-clustered condition). The location of words in the non-clustered condition are not relevant to the "human-made" pattern in this case, so we define these items as "schema-irrelevant".

Though evidence suggests that schemas can bias memory by increasing false alarms (Neuschatz et al., 2002) and increasing the number of false memories (Kleider et al., 2008), it is not often considered how schemas influence information that is not relevant to themselves. Though some studies have included irrelevant information in their paradigm (e.g., Frank et al., 2018; Greve et al., 2019), this was used as a control condition to compare performance relative to congruent and incongruent information, as opposed to examining how the presence of schematic information could bias behaviour for this irrelevant information. Indeed, our schema-irrelevant (non-clustered) condition was first created as a control condition before we focussed our attention subsequently on behavioural biases specifically in this condition.

Returning to precision measures of memory and generalisation, Tompary et al. (2020) did not include a control condition where locations for one semantic group were randomly distributed. Instead, they used two clustered distributions separated by 180^{o}, so it is difficult to disentangle the effects of one cluster against another. In the present experiments, we used the clustered and non-clustered conditions introduced in Berens et al. (2020) and introduced novel semantically related items (as in Tompary et al., 2020). This allowed us to focus on behaviour in the non-clustered condition, where the words are from a separate semantic category to the clustered condition, and the locations of these words are randomly distributed. As such, word-locations in the non-clustered condition are technically irrelevant to extracting the underlying pattern (or schema) in the clustered condition.

2.2.1. Overview of Experiments

We explored how the presence of a pattern influences memory and generalisation when one condition possesses a pattern and the other does not. We used an experimental design similar to Berens et al. (2020), but with the inclusion of novel items at test. Participants learned word-location associations around a circle. Word stimuli came from two semantic categories: human-made (e.g., chair, computer) and natural (e.g., leaf, giraffe). The locations associated with these words were either clustered or non-clustered. By including the non-clustered condition, we could explore how the presence of a pattern affected behaviour for information semantically related to the pattern (i.e., words belonging to the clustered category) and semantically unrelated to the pattern (i.e., words belonging to the non-clustered category). Using the generalisation trials, we can assess the impact of schematic information on behaviour directly without the noise added by memory trials. Whilst memory trials can rely on both memory for the individual item and the schematic information; generalisation trials may only rely on the schematic information or random guessing. Therefore, it provided an opportunity to establish schematic influence on irrelevant information. Specifically, participants may form a 'schematic' representation for the semantic category associated with the clustered condition, allowing them to make predictions about the possible locations of novel words belonging to the same category. In contrast, for the semantic category associated with the non-clustered condition, there was no underlying pattern. This allowed us to observe how a pattern in the clustered condition influences generalisation behaviour for 'schema-irrelevant' words. Across four experiments, we manipulated delay between Study and Test, and whether we collected data in person (in the lab) or online, providing evidence that schemas bias memory and generalisation behaviour in the schema-irrelevant (non-clustered) condition.

2.3. Experiment 1

In Experiment 1, we asked two questions: (1) does the presence of a pattern increase memory performance in the clustered relative to non-clustered condition, and (2) can participants generalise, such that they place novel words in locations similar to the pattern in the clustered relative to the non-clustered condition? To answer these questions, we had two pre-registered hypotheses: (1) participants' overall memory performance will be greater in the clustered relative to non-clustered condition (as measured by 'Total Information', see Methods), and (2) the distribution of locations for novel words will be more similar to the underlying pattern (von Mises distribution) in the clustered relative to non-clustered condition (whereas the distribution in non-clustered condition will be more uniform; as measured by Kullback-Leibler divergence). The preregistration for Experiment 1 is available at: https://osf.io/h6wba/.

2.3.1. Methods

2.3.1.1. Participants

2.3.1.1.1. Power Analysis

Two power calculations were conducted to estimate the required sample size to examine the pre-registered hypotheses. First, to estimate the required sample size for the effect of clustering on total information, G*Power (3.1.9.2; Faul et al., 2007) was used to perform an *a priori* power analysis. A power analysis was computed for a paired samples *t*-test comparing total information in the clustered and non-clustered conditions. The effect size for the analysis was estimated from a pilot investigation reported in Berens et al. (2020). This pilot study estimated an effect of d = 0.33, with the clustered condition showing significantly greater total information than the non-clustered. This effect size estimate, along with an α

(one-tailed) = .05 and power = .80 were used. A suggested sample size of N = 59 usable datasets was estimated.

Second, data simulations were conducted to estimate the required sample size to compare the distribution of locations for clustered and non-clustered novel words to the underlying experimental distribution of clustered items (i.e., von Mises distribution). Data simulations were run to identify: (1) the minimum number of responses required to get a reliable estimate of Kullback-Leibler divergence (D_{KL}), and (2) to determine the required sample size to gain 80% power. The number of participants and words varied on each iteration of 100 simulations. The simulation assumed that each participant reproduced the spatial distribution of clustered and non-clustered locations with varying accuracy. Specifically, the reproduced distributions took the form of a von Mises probability density with a mean parameter drawn from a von Mises (μ = 0, κ = 5.5). The concentration parameter for each distribution was then sampled independently from a gamma distribution with a mean of 2 (i.e., the true concentration) and a standard deviation of 2. These parameters were estimated from a previous pilot study by Berens et al. (2020). Non-parametric density functions were then estimated from the simulated responses in both conditions separately. The probability density across circular locations was then compared to the experimentally imposed von Mises distribution (μ = 0 and κ = 2) using D_{KL}. A generalised linear mixed-effects (GLME) model, using the same parameters as described below (see Data Handling), was fit to the D_{KL} measures of both clustered and non-clustered responses with varying intercepts based on each 'participant'. No random slopes were computed for these simulations. It was found that a minimum of 11 words and 9 participants were required. The code used to generate these simulations can be found on the OSF page (https://osf.io/bxru4/). Given the above, a final sample of 60 usable datasets was pre-registered.

2.3.1.1.2. Final Sample

Sixty-nine participants (63 female) were recruited for the study. The mean age was 19.59 years (*SD* = 2.22 years). The mixture model failed to converge for 6 participants, meaning the final sample consisted of 63 participants (58 female) with a mean age of 19.63 years (SD = 2.30 years). The over-recruitment resulted from a minor coding error resulting in the incorrect rejection of valid model fits for three participants. Participants were fluent English-speakers with normal or corrected-to-normal vision and were recruited from the University of York student population and took part in exchange for course credit. Ethical approval for all experiments was granted by the Department of Psychology Ethics Committee at the University of York. Exclusion criteria for the data are detailed below (see Exclusion Criteria).

2.3.2. Materials: Word Lists

To develop the word lists, semantic representations of 324 words were extracted from a pre-trained word2vec model (Mikolov et al., 2013). This pre-trained model of numerical word representations contained over 3 million English words based on the Google News dataset. The semantic similarity between word representations was then computed via Euclidean distance. Simulations were run to ensure a small semantic distance between words of the same category (e.g., natural) whilst ensuring a large semantic distance between cross-category pairs. To meet these criteria, simulations were run using 10,000 iterations to identify a word list containing a total of 240 (120 human-made and 120 natural) words. The final list had a mean semantic distance of 4.24 (*SD* = 0.47) within and 4.44 (*SD* = 0.42) between categories, suggesting the two lists were sufficiently distinct in terms of semantic grouping. The distributions of semantic distances within each group were comparable as compared using the Kolmogorov-Smirnov test (*D* = 0.01). After generating the lists, we ensured word length and frequency of use in natural language, as quantified using the Zipf-scale of the

SUBTLEX-UK database (van Heuven et al., 2014), were comparable across lists. Finally, to split the two lists into eight sets of 30, a further 10,000 iterations were run. Sub-lists were generated by controlling for the mean and variance in Euclidean distance, the distribution difference using the Kolmogorov–Smirnov test, word frequency and word length. The code and word lists used can be found here: https://osf.io/bxru4/.

2.3.3. Procedure

2.3.3.1. Study Phase

Participants learned associations between different locations around a circle and a specific word displayed on each trial. During the study phase, 180 words were presented. One of the semantic categories was assigned to the clustered condition (counterbalanced across participants). Word-locations in this condition were clustered by sampling from a von Mises distribution with a fixed width (κ = 2.0) and a fixed mean (randomly selected for each participant). The other semantic category was assigned to the non-clustered condition. Word-locations in this condition were randomly distributed around the circle by sampling from a uniform distribution. Participants were not informed about the presence of the semantic categories or the clustering manipulation. They were only told that they would need to remember each individual word-location association.

All stimuli were presented using MATLAB and the COGENT 2000 toolbox (www.vislab.ucl.ac.uk/cogent/index.html) on a desktop PC. Participants sat approximately 50cm away from the screen so that the circle subtended ~16 visual degrees. Each study trial (shown in Figure 2.1) started with a fixation cross for 1s, followed by a location marker, which was present for 2s. The location marker and circle were then removed and the study word displayed for 4s. Subsequently, with the word still present, the circle and marker, the latter of which was redrawn at a random location around the circle, were presented. Participants were asked to use a mouse to reposition the cursor back to the cued location; this response window lasted 6s. Repositioning the marker during study ensured participants deliberately attended to the word-location association as opposed to passively viewing. If participants did not respond within the 6s time window or selected an area greater than 10° from the presented location, the trial was repeated, with a red fixation cross at the beginning of the trial to alert them to this repetition. The average number of repetitions across all experiments reported in this Chapter was 0.17 trials (*SD* = 0.45, Proportion = 0.002) and 0.15 trials (*SD* = 0.43, Proportion = 0.002) for the non-clustered and clustered conditions, respectively.

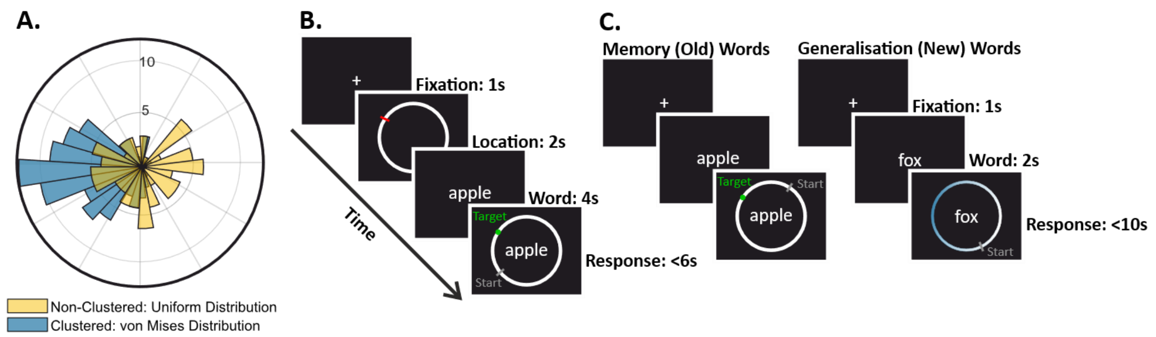

Figure 2.1. **Precision Paradigm: Experimental Design**. (A) **Clustering**: The polar plot shows an example of distributed locations for one participant. The clustered and non-clustered conditions were associated with either the human-made or natural word category (counterbalanced across participants) and the centre of the clustered distribution was randomised for each participant. Numbers on the polar plot show the number of words located in that area of the circle. (B) **Study Phase**: Participants were presented with a fixation cross (1s), followed by the location (2s), then the word alone (4s), and then presented with the word, the circle, and a randomly placed marker to make a response (6s). Participants moved the marker from the start location back to the location just presented. (C) **Test Phase**: Participants were first present with a fixation cross (1s), the word alone (2s) and then asked to replace the marker from the randomly generated start position back to the remembered location (memory trials) or make an inference based on experience (generalisation trials, 10s). In the example above, natural words were assigned to the clustered condition. The blue shading in the generalisation trial shows the area of the circle they are likely to generalise to in the clustered condition.

Before starting the study phase, participants were given practise trials to ensure they understood the task and knew how to make responses. The practise trials used similar parameters as described above, but with abstract nouns (e.g., beauty, jealousy, integrity) that held no semantic clustering and no relation to words within the study lists. There was a total of 10 practise trials. Following the Study phase, participants took part in an immediate Test phase.

2.3.3.2. Test Phase

At Test, participants were required to recall the 180 previously presented word-location associations and select locations for novel words (60 words). These novel words came from the same semantic groupings as above. The old and new words were intermixed, and presentation order was randomised. On each test trial (Figure 2.1), a fixation cross appeared for 1s, followed by the presentation of the word for 2s before the circle and marker appeared, with the marker being presented at a random location around the circle. Using a mouse, participants had 10s to reposition the marker back to the remembered location, or to make a best guess if they had forgotten. Participants were not told about the presence of novel words at Test, with the trial structure being identical. Participants were told to make a best guess for any words where they had forgotten the location.

2.3.3.3. Introspection Questionnaire

Following Test, participants completed an Introspection Questionnaire. The questionnaire addressed their perceptions on task difficulty, asked them to report their strategies for words they had forgotten, whether they noticed any words presented at Test that were not presented at Study, their strategies for these words, and whether they felt a pattern was present in the presentation of word-location associations. The questionnaire is located here: https://osf.io/7fgzm/.

2.3.4. Data Handling

2.3.4.1. Mixture Model Estimation

Using mixture modelling, we estimated accessibility (i.e., word-location retrieval probability) and precision (i.e., how precisely are locations remembered given they are accessible) for individual participants. We calculated the replacement error for each response (i.e., the angular difference between the correct location and remembered location). These angular errors are assumed to come from one of two distributions: (1) a circular uniform distribution representing guesses, and (2) a von Mises distribution representing accessible word-location associations, whose variance represents the degree of 'precision' that locations were remembered. These two distributions have associated prior probabilities, which reflect the overall proportion of responses belonging to either distribution. For the von Mises distribution, the prior (p) represents retrieval probability (i.e., accessibility). This distribution also has two other parameters: mean (μ) and concentration (κ). The value of μ was fixed at zero, assuming the average error of responses was zero. The κ represents the variance, or precision, in responses. Higher κ values indicate a narrower distribution (higher precision), lower κ values indicate a wider distribution (lower precision).

Mixture modelling was conducted using the HoopStats toolbox developed in Berens et al. (2020), found here: https://osf.io/8mzyc/. First, an Expectation Maximisation (EM) algorithm was used to estimate p and κ for each participant, and clustered and non-clustered items, separately. The overall fit of this model was then compared to a reduced model where all angular errors are assumed to be from a uniform distribution (i.e., no mnemonic information is present). This comparison was conducted using the Bayesian Information Criterion (BIC). If the BIC was less than -10 (i.e., evidence in favour of the two-distribution model), the parameters returned from the EM were accepted. If, however, the BIC was greater than -10, representing a poorly fit model, an alternative fitting procedure was implemented. This failure to meet criterion often occurs when low accessibility is present in the data ($p \lesssim .2$). For the alternative fitting procedure, the p value was systematically varied over several steps for this alternative model, with κ being estimated from the corresponding responses with the smallest angular error. Using this method, valid model fits could be found that were otherwise missed by the EM algorithm. If this alternate model produces a better fit than the single uniform distribution, again using the BIC < -10 criterion, these parameters were accepted. If BIC > -10, or the estimates of κ were modelled on fewer than eight trials, the participant's entire dataset was excluded.

2.3.4.1.1. Conversion to Entropy Measures

Once both the p and κ parameters were estimated for clustered and non-clustered trials (as in Berens et al., 2020), both p and κ were converted into information entropy measures I_p and I_κ. Small values of I_p indicate lower levels of accessibility. Similarly, small values of I_κ indicate poor precision. Conversion of p and κ to I_p and I_κ allows for a more direct comparison, as they describe performance using the same metric: information gain (in nats) relative to random responses. Additionally, we computed a combined measure of memory performance, "Total Information" (I_t), which is directly proportional to both I_p and I_κ ($I_t = \frac{Ip*Ik}{\log{(2\pi)}}$). I_t reflects the total amount of mnemonic information present at the point of retrieval, which is a function of both the proportion of word-location pairs that were accessible and the precision of these accessible word locations. Hypothesis 1 uses this measure of Total Information to assess overall memory performance between the clustered and non-clustered conditions.

2.3.4.2. Kernel Density Estimation

Kernel density estimates were computed to characterise the distribution of location responses; this was identical to Berens et al. (2020). The primary purpose of the kernel density estimates was to compute the Kullback-Leibler divergence (D_{KL}) between participants' responses and the pattern of studied locations. They were also used: (1) to plot the distribution of angular errors for memory trials, and (2) to plot the distribution of responses relative to the experimentally imposed von Mises distribution for memory and generalisation trials. To do this, a von Mises probability density function, with a concentration of $\kappa = 2$, was centred on each response. This distribution acted as a smoothing kernel that spread a small portion of the overall density around the local area. As such, the density estimates at a given angle were taken as the mean probability density value across all these distributions. The responses were either angular errors for each condition (for memory trials) or angular differences between the responses and the centre of the experimentally imposed cluster (for generalisation trials).

2.3.4.3. Kullback-Leibler Divergence

Once the spatial distribution of responses was estimated through the kernel density function, D_{KL} was computed to assess the similarity between specific distributions. D_{KL} measures divergence between two distributions, with higher values representing greater divergence (i.e., less similarity) between the two; this was computed via numerical integration, as in Berens et al. (2020), rather than using a discrete approximation. First, we assessed how divergent the distributions for clustered and non-clustered novel words (i.e., generalisation trials) were to the reference distribution (i.e., the experimentally imposed von Mises distribution associated with the clustered condition). The distribution of clustered novel words was predicted to be less divergent to the underlying von Mises pattern relative to the non-clustered condition (Hypothesis 2). Second, we assessed how divergent the distributions for clustered and non-clustered novel words were to a uniform distribution (i.e., no pattern). The distribution of non-clustered novel words was predicted to be less divergent from a uniform distribution relative to the clustered condition (Hypothesis 2).

2.3.4.4. Exclusion Criteria

All exclusion criteria were pre-registered. If an additional exclusion was included that was not pre-registered, this is explicitly identified throughout the chapter.

At the trial level, individual trials would be excluded for two reasons. First, if participants repeated trials during the Study Phase five times or more due to a lack of responding or being outside the 10° limit, it was removed from later statistical analyses to ensure that the extra encoding of these word-location pairs did not impact retrieval. This cut-off was selected based on an observation made during the in-lab piloting for Berens et al. (2020), where very few participants needed to repeat a trial on more than five occasions, with only ~10% of participants requiring greater than five repetitions to replace the marker within 5° of the presented location. For all four experiments in this chapter, only 13 trials across participants and experiments were repeated more than five times showing few trials were removed for this reason. Trials were also removed if no response was given at Test to ensure only trials where participants gave an explicit response were included. Across all experiments reported, on average, participants did not respond to 1.75 trials (SD = 4.23).

Datasets would only be included for analysis when the following criteria were met: (1) both the study and test trials were complete, (2) the number of old words with no response did not exceed 20 trials for the clustered and non-clustered conditions separately, (3) the number of novel words not responded to did not exceed 15 trials for the clustered and non-clustered conditions separately, (4) the dataset was not corrupted, and (5) the mixture model could be fit adequately to the data (see: Mixture Model Estimation, above).

2.3.5. Statistical Analysis

All statistical analyses reported in the main results sections of all experiments were pre-registered. Where exploratory analyses were run, these are clearly labelled as such. We performed three separate GLME models. The models were used to predict (1) Total Information (I_t), (2) D_{KL} in comparison to the experimentally imposed von Mises distribution, and (3) D_{KL} in comparison to a uniform distribution. The first model relates to Hypothesis 1, assessing whether overall memory performance differs between the clustered and non-clustered conditions. The second and third models relate to Hypothesis 2, testing whether participants can position novel words from the same semantic category according to the underlying pattern.

For each model, we compared the clustered and non-clustered conditions for each measure of interest. All models were fit to the data using a log link function and gamma distribution to model the spread of the data. The model was estimated using the maximum likelihood fitting procedure in the MATLAB Statistics and Machine Learning Toolbox. The models included the independent variable of Clustering (0 = Non-Clustered; 1 = Clustered). In addition to this fixed effect, a set of random effect parameters (two per participant) were included. One random effect allowed the intercepts to vary based on participant, and the other allowed the effect of clustering to vary by participant. All elements of the associated random effects covariance matrix were estimated from the data. The model did not converge for the D_{KL} model that compared clustered and non-clustered novel locations to the uniform distribution. As a result, the random slopes for clustering were removed for this comparison across experiments where this analysis is reported.

All mean values represent the mean estimate of the population derived from the GLME. Further, the Cohen's *d* and Bayes Factors (BF_{01}) reported were calculated as reported in Berens et al. (2020) and estimated only on the fixed factors. In brief, the model parameters were used to generate effect size estimates, which were then used to calculate BF_{01} evidence in favour of the null model. We pre-registered that Bayes Factors would only be reported for non-significant results to aid in interpreting the outcome of these tests by assessing whether there was greater support for the null relative to the alternative hypothesis. However, we feel these can be informative for both significant and non-significant results. A further deviation from the pre-registration was how Bayes Factors were computed. Previously we specified Bayes Factors would be computed in JASP. However, to increase reproducibility, the computation used in Berens et al. (2020) was used for all Bayes Factors reported. For these analyses, a prior Cauchy distribution of r = .707 centred at 0 was used; this was identical to the pre-registration. All analyses use two-tailed tests unless otherwise specified.

2.3.6. Results

2.3.6.1. Memory

Figure 2.2 shows the Total Information metric for Experiments 1-4 and the probability density estimates for angular error. The angular error plots demonstrate differences, across the delay periods, in the degrees of error around the circle. These plots demonstrate possible differences between conditions based on accessibility and precision. Specifically, the higher peaks in the clustered condition suggest greater accessibility, whilst the narrower distributions for the non-clustered condition suggest greater precision. These metrics are analysed in the Across Experimental Analysis section towards the end of the Chapter.

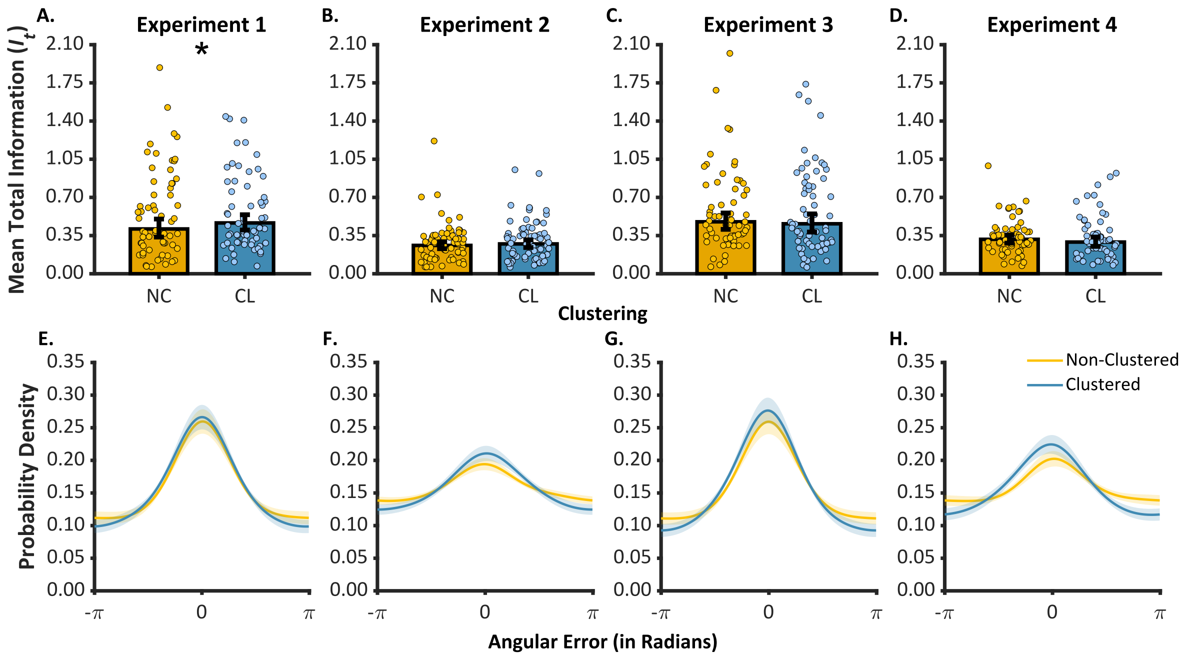

Figure 2.2. **Overall Memory performance across experiments**. *A-D*: Mean Total Information (I_t) across Experiments 1-4 as a function of clustering (clustered and non-clustered). Individual data points represent participant scores. *E-G*: Spatial distribution of angular errors across experiments, 0 here represents 0° of error. Error bars represent 95% confidence intervals around the mean for all plots. * = $p < .05$. CL = Clustered. NC = Non-Clustered.

Hypothesis 1 related to whether clustering benefits overall memory performance. Consistent with this, in Experiment 1, total information was significantly greater in the clustered relative to the non-clustered condition, $t(124) = 1.99$, $p = .049$, $d = 0.35$, $BF_{01} = 0.87$. Though significant, the Bayes Factor was inconclusive.

2.3.6.2. Generalisation

Figure 2.3 shows generalisation behaviour for the novel words for Experiments 1-4. Hypothesis 2 was that the distribution of selected locations for clustered generalisation trials would be more similar to (i.e., less divergent from) the experimentally imposed von Mises distribution than the distribution of locations for non-clustered generalisation trials. This was corroborated statistically, where the clustered responses were found to be significantly less divergent from the von Mises distribution than the non-clustered responses, $t(124) = 4.26$, $p < .001$, $d = 0.76$, $BF_{01} = 0.001$. This suggests participants could make reasonable guesses or predictions about where novel words would be located based on the learnt locations from the same semantic category. Specifically, participants placed novel words in the clustered category in similar locations to the old clustered items relative to novel words in the non-clustered category.

We also predicted that the distribution of locations for non-clustered novel words would be more similar to (less divergent from) a uniform distribution relative to the distribution for clustered words. In other words, we expected greater uniformity (or 'entropy') of responses in the non-clustered relative to the clustered condition. Inconsistent with this prediction, no difference in D_{KL} was observed, $t(124) = 0.08$, $p = .933$, $d = 0.02$, $BF_{01} = 5.24$. Indeed, Bayes Factors indicated there was five times more evidence in favour of the null, suggesting the clustered and non-clustered conditions diverged equally from the uniform distribution. Interestingly, the kernel density estimates at the centre of the experimentally imposed distribution ($\vartheta = 0$) show an increase for clustered responses but a decrease for non-clustered responses (Figure 2.3E). Thus, despite the two conditions having equally diverged from the uniform distribution, they may have diverged in a qualitatively distinct manner. We return to this finding following Experiment 2.

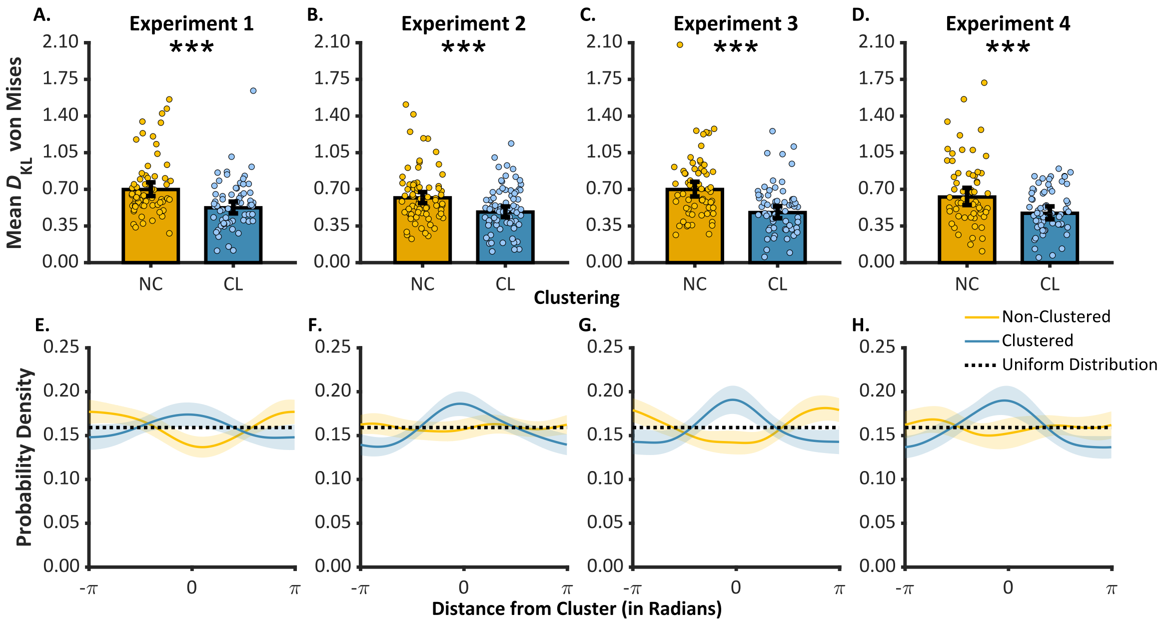

Figure 2.3. **Generalisation behaviour across experiments**. *A-D*: Mean divergence (D_{KL}) from the experimentally imposed von Mises distribution across Experiments 1-4 as a function of clustering (clustered and non-clustered). Individual data points represent participant scores. *E-G*: Spatial distribution of locations selected for novel words, centred to the experimentally imposed von Mises distribution. Error bars represent 95% confidence intervals around the mean for all plots. *** = $p \leq .001$. CL = Clustered. NC = Non-Clustered.

2.3.7. Discussion

Experiment 1 assessed how the presence of an underlying pattern (or schema) modulated memory and generalisation behaviour. Memory performance was greater for the clustered relative to the non-clustered condition (Hypothesis 1). Additionally, when presented with novel words, participants reproduced the pattern of locations presented for the clustered items, meaning they showed an ability to generalise their mnemonic information to novel, semantically related, words (Hypothesis 2). However, both conditions were equally divergent from the uniform distribution, which was not in line with expectations.

The finding that memory was benefited by the presence of a pattern is consistent with previous studies (Atienza et al., 2011; Brewer & Treyens, 1981; Greve et al., 2019). However, we note that Berens et al. (2020) did not find a difference in total information between the clustered and non-clustered conditions. They did see differences in accessibility and precision, a finding we will return to later.

Generalisation of the clustered items was found to be more similar to the experimentally imposed pattern than for the non-clustered items. These findings are consistent with recent evidence showing generalisation to novel instances can occur rapidly without the need for an extended period of consolidation (e.g., Sweegers & Talamini, 2014; Zeithamova et al., 2012).

Interestingly, the distribution of locations for clustered and non-clustered novel words diverged equally from a uniform distribution. Inspection of Figure 2.3E suggests participants may have been less likely to place novel words in the non-clustered condition near the centre of the clustered distribution; a possible "avoidance" effect. This may suggest that the presence of a pattern (i.e., schema) in one condition influences schema-irrelevant information in the non-clustered condition. We return to this following Experiment 2.

An open question was whether generalisation behaviour was modulated by delay. Theories of systems consolidation suggest that the extraction of schemas across a set of related experiences may take time to emerge (Kumaran & McClelland, 2012; McClelland et al., 1995), and sleep may play a critical role in this process (Inostroza & Born, 2013). Behavioural (Tompary et al., 2020) and neuroimaging (Kroes & Fernández, 2012; Wagner et al., 2015) work also suggests a time-dependent effect either in terms of the use or establishment of a schema. As such, a delay between Study and Test may allow us to see more evidence of generalisation compared to an immediate Test phase. Therefore, Experiment 2 sought to replicate Experiment 1 with one change – adding a delay between Study and Test.

2.4. Experiment 2

Experiment 2 was identical to Experiment 1 with one exception – we increased the delay between Study and Test to approximately 24-hours. The same preregistered hypotheses from Experiment 1 were tested. The preregistration for Experiment 2 is available here: https://osf.io/nbtm3/.

2.4.1. Methods

2.4.1.1. Participants

2.4.1.1.1. Power Analysis

To determine the required sample size, the smallest effect size of interest for the pre-registered hypotheses was used; this was derived from the Berens et al. (2020) pilot investigation and concerned the effect of clustering on total information following a 24-hour delay between Study and Test. As before, G*Power (3.1.9.2; Faul et al., 2007) was used to perform an *a priori* power analysis for a paired samples *t*-test comparing total information in the clustered and non-clustered conditions. The effect size for the analysis was estimated from the pilot investigation by Berens et al. (2020), which derived an effect size of d = 0.31. This effect size estimate, along with an α (one-tailed) = .05 and power = .80 were used. A suggested sample size of N = 66 was required.

2.4.1.1.2. Final Sample

Eighty-six participants (77 female) were recruited for the study. The mean age was 20.63 years (SD = 2.01 years). Three participants did not return for the second session, and eight datasets did not converge using the Mixture Model and so were excluded. The final sample consisted of 75 participants (67 female) with a mean age of 20.55 years (SD = 1.99 years). Similar to Experiment 1, when first analysed, a lower number (66 participants) of usable datasets were present. However, when a coding issue was fixed, a sample of 75 usable datasets was obtained (hence the over-recruitment). Participants were fluent English-speakers with normal or corrected-to-normal vision and were recruited from the University of York student population and took part in exchange for course credit or cash payment.

2.4.2. Materials and Procedure

The same materials and procedure were followed from Experiment 1; however, participants completed the Test phase approximately 24-hours post Study, with the average delay between study and test being 23.93 hours (*SD* = 0.39 hours).

2.4.3. Data Handling and Statistical Analysis

The same exclusion criteria and statistical analyses were used as in Experiment 1.

2.4.4. Results

2.4.4.1. Memory

When assessing memory performance, unlike Experiment 1, total information did not significantly differ between the two conditions, $t(148) = 0.67$, $p = .502$, $d = 0.11$, $BF_{01} = 4.63$ (Figure 2.2B; Hypothesis 1). The Bayes Factor indicates four times more support favouring the null model, suggesting no difference between conditions was present.

2.4.4.2. Generalisation

Figure 2.3 suggests a similar pattern of results to Experiment 1. The clustered condition was significantly less divergent from the von Mises distribution than the non-clustered condition, $t(148) = 3.29$, $p = .001$, $d = 0.54$, $BF_{01} = 0.04$. In comparison to the uniform distribution, neither condition was significantly more divergent than the other, $t(148) = 0.44$, $p = .663$, $d = 0.07$, $BF_{01} = 5.22$. These results replicate Experiment 1.

2.4.4.3. Exploratory Comparison of Experiments 1 and 2

2.4.4.3.1. Change in Generalisation

Previous work suggests that schema may take time to develop, with a period of sleep being an important contributor to this development (Inostroza & Born, 2013). As such, we wished to assess whether generalisation behaviour for the clustered novel items changed following a delay period. We predicted that generalisation would be greater (represented by lower D_{KL} values) following a delay. Figures 2.3A and 2.3B show the mean divergence for both conditions across experiments.

To assess a change over time, we computed a GLME using the same parameters for the previous GLME's reported above. However, instead of the effect of clustering, we assessed the effect of Delay (0 = Immediate Test, 1 = Delayed Test) on the divergence between the experimentally imposed von Mises distribution and responses for clustered novel words. Further, no random slopes were included in the model, with only random intercepts for each participant. It was found that there was no significant evidence of a change over time ($t(136) = 1.08$, $p = .280$, $d = 0.19$, $BF_{01} = 3.20$). This suggests that following a period of consolidation, participants' adherence to the von Mises distribution for clustered novel items did not change.

2.4.4.3.2. Avoidance Behaviour

In Figures 2.3E and 2.3F, there was possible evidence for a lack of uniformity in the distribution of locations for novel words in the non-clustered condition. Participants appear to avoid the centre of the cluster for novel non-clustered words (though visual inspection suggests this effect is perhaps greater in Experiment 1 than 2). To assess this possible avoidance more formally, we compared non-clustered kernel density estimates at the centre of the cluster to the density expected if the responses were uniformly distributed ($2\pi^{-1}$). If participants were actively avoiding the centre of the cluster, their kernel density for non-clustered items at this location would be significantly lower than the uniform value. A GLME was computed using the same parameters described for the model change in generalisation model above.

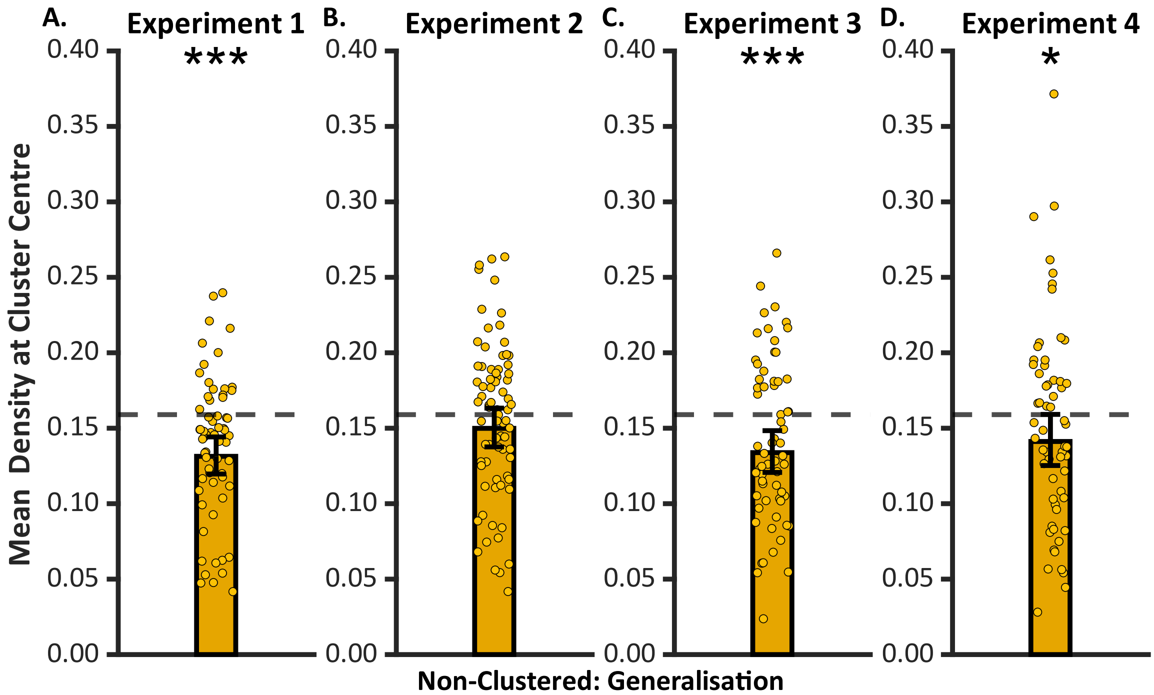

Figure 2.4. **Avoidance in generalisation across experiments**. Experiments 1, 3 and 4 all show evidence of an avoidance effect, with significantly lower kernel density at the centre of the cluster. Error bars are 95% confidence intervals around the mean estimate. Individual data points represent participant scores. The dashed line represents the uniform probability value (0.159). * = $p < .050$, *** = $p < .001$.

The kernel density plots for all experiments are shown in Figure 2.4. For the analysis, we first compared Immediate Test and Delayed Test to the uniform value, separately. At immediate test, there was significantly reduced probability density compared to the uniform value, $t(136) = 4.06$, $p < .001$, $d = 0.35$, $BF_{01} = 0.003$; this suggests participants actively avoided locations at the centre of the cluster for novel non-clustered items. In contrast, there was no significant reduction in probability density during Delayed Test, $t(136) = 1.38$, $p = .169$, $d = 0.10$, $BF_{01} = 4.65$. Along with this, a significant effect of delay was observed, $t(136) = 2.06$, $p = .041$, $d = 0.35$, $BF_{01} = 0.78$. Here, there was a decrease in the avoidance effect in Experiment 2 relative to Experiment 1 (i.e., distributions of novel words were more uniform following a delay). However, the Bayes Factor was anecdotal.

2.4.5. Discussion

Experiment 2 replicated the key generalisation finding from Experiment 1 – participants' distributions of locations were more similar to the underlying pattern in the clustered condition than the non-clustered condition (Hypothesis 2). However, we did not replicate the difference in overall memory performance (Hypothesis 1). The lack of difference in total information instead agrees with the results of Berens et al. (2020). Additionally, we found that, following a delay period, participants' adherence to the underlying pattern did not change for clustered items. This is contrary to some lines of evidence suggesting schematic extraction may take time to develop (e.g., Inostroza & Born, 2013; Kumaran & McClelland, 2012; McClelland et al., 1995), though other studies report that generalisation based on an underlying pattern remains relatively stable over extended periods (e.g., 1-month post-learning, Sweegers & Talamini, 2014).

We also saw an "avoidance effect" in the non-clustered condition, where participants avoided placing novel words in the non-clustered condition at the centre of the cluster. Despite old non-clustered words being drawn from a uniform distribution and being from a separate semantic category to the clustered words, participants were biased away from the clustered location. This avoidance effect was present in Experiment 1 and decreased in Experiment 2, where it was no longer present. Thus, this avoidance effect appears immediately but possibly decreases over a 24-hour delay (though see results of Experiments 3 and 4). Given that this avoidance effect was not predicted, we performed two further experiments with pre-registered analyses to replicate this effect.

2.5. Experiment 3

Two further experiments were conducted. Experiment 3 aimed to replicate Experiment 1 with the Test phase immediately following the Study phase. Experiment 4 aimed to replicate Experiment 2, with a 24-hour delay between Study and Test. Both experiments were run online due to coronavirus restrictions. The hypotheses from Experiments 1 and 2 were repeated in Experiments 3 and 4. However, the comparison of the generalisation trials distributions to a uniform distribution was excluded as this comparison was not informative in Experiments 1 and 2. Critically an additional preregistered analysis was included concerning the avoidance effect in the non-clustered generalisation condition; this was the same as the exploratory analysis of Experiments 1 and 2 (see Statistical Analysis, below). We predicted that participants would show a significant reduction in probability density for non-clustered novel words at the centre of the cluster (relative to a uniform distribution, as in Experiment 1). The preregistration for Experiment 3 is available here: https://osf.io/2wsn8/.

2.5.1. Methods

2.5.1.1. Participants

2.5.1.1.1. Power Analysis

To determine the required sample, we assessed the range of effect sizes from Experiment 1 (d = 0.35 – 0.77) and set a minimum effect size of theoretical interest (i.e., Hypothesis 3, d = 0.35). A power analysis of a one-sample t-test with the effect size of interest, α = .05 (one-tailed) and power = .80 was conducted using G*Power (3.1.9, Faul et al., 2007). A sample size of 52 usable datasets was needed. However, given this estimate, and the power analysis previously conducted for Experiment 1, a final sample size of 60 usable datasets was set.

2.5.1.1.2. Final Sample

Eighty-nine participants (35 female) with a mean age of 24.91 years (*SD* = 4.99 years) were recruited for the experiment. Three participants left before the study phase, ten were excluded during Study, one left before completing the Test phase, three were excluded at Test due to inattention, and two attempted the study phase twice and so were excluded. This left 70 participants that passed the initial checks. Of those, four did not respond to 20 or more memory trials, and three datasets did not converge during the mixture model. Therefore, the final sample was 63 participants (25 female) with a mean age of 25.27 years (*SD* = 4.99 years). All participants were fluent English-speakers with normal or corrected-to-normal vision and were recruited through Prolific.co and received monetary compensation for their time.

2.5.1.2. Materials

The same word lists and Introspection Questionnaire were used from Experiments 1 and 2. However, rather than using four sub-lists for each category (i.e., human-made and natural) as in Experiments 1 and 2, 30 words from each category were randomly selected for each participant and assigned to the generalisation condition. The remaining 90 words were assigned to the memory condition. This was done due to practical constraints when coding the online experiment.

2.5.1.3. Procedure

The same general procedure was followed as in Experiment 1, but through the online platform Prolific. Participants recruited from Prolific were directed to a secure website hosting the online experiment. Participants could only use a laptop or desktop computer to run the task, with handheld devices (e.g., smartphone, tablet) being excluded. Before starting the Study Phase, participants watched a short introductory video about how the session progressed and how to respond. A PDF document of written instructions was also provided (https://osf.io/qxfuj/). The instructions emphasised the need to visualise the object related to the cue word appearing at the cued location before responding on each study trial and how participants were to be asked to recall these locations at test. The video instructions replaced the practise trials used in-lab, as using instructions in this format online produced similar results for memory trials in the Berens et al. (2020) study.

2.5.1.3.1. Study Phase

The Study Phase was identical to Experiments 1 and 2. Once completed, participants moved immediately onto the Test Phase.

2.5.1.3.2. Test Phase

One minor change was made to the Test phase. In Experiments 1 and 2, participants were presented with a fixation cross (1s) followed by the word alone (2s) and then the opportunity to reposition the marker to the remembered or generalised location (10s). In Experiments 3-4, the word was not shown alone for 2s. In-lab, participants provided a response on average within 2.38s of being able to replace the marker, with almost all responses collected within 7.23s. As such, the additional 2s of the word alone was removed given the 10s-time window for responding. Following Test, participants were asked to complete the Introspection Questionnaire.

2.5.1.3.3. Introspection Questionnaire

The same questions as Experiments 1 and 2 were used online. We also included an additional question about whether the participant had help completing the task; this was to be used as an exclusion criterion (though not pre-registered) had participants reported they did have help completing the task. No such report was given.

2.5.1.4. Data Handling and Statistical Analysis

2.5.1.4.1. Exclusion Criteria

All exclusions from Experiments 1 and 2 were used in Experiments 3 and 4. However, participants could also be excluded during Study or Test for not following task instructions; this was quantified as having reaction times of less than 2s across a total of 70 trials. Specifically, participants would receive a warning message through their browser should the number of trials with reaction times less than 2s hit 10, 30, 45 and 60 trials. This message asked participants to either: slow down and ensure they imagined the object appearing at each location (Study Phase) or encouraged them to remember the location for each word (Test Phase). This was an exclusion that was not pre-registered but used previously (see Berens et al., 2020) as a way of maximising participant performance when experimenting online.

2.5.1.4.2. Statistical Analysis

The same statistical analyses were used for Hypotheses 1 (Total Information) and 2 (D_{KL} von Mises) as described previously. For Hypothesis 3, we compared the probability density estimates at the centre of participants experimentally imposed cluster for non-clustered novel words to the density for a uniform distribution. If participants were actively avoiding the centre of the cluster, then they will not be distributing locations randomly (or uniformly), so their kernel density at this location should be significantly below that of a uniform value. To test this, a GLME was fit using a log link function and a gamma distribution to model the spread of the data, estimated using the maximum likelihood estimate fitting method within the MATLAB Statistics and Machine Learning Toolbox. This was an intercept only model with random intercepts for each participant. The derived model was then used to conduct a one-sample *t*-test comparing the beta of the intercept model to the log of the uniform kernel density value. A one-tailed test was used for this analysis as a directional effect was predicted. Note, this analysis was almost identical to the exploratory analysis performed across Experiments 1 and 2 but without the inclusion of any fixed effects. Cohen's *d* and Bayes Factors are reported and use the same parameters as described previously.

2.5.2. Results

2.5.2.1. Memory

Total Information was not significantly different between the clustered and non-clustered conditions, $t(124) = 0.73$, $p = .466$, $d = 0.13$, $BF_{01} = 4.13$ (Figure 2.2C). There was four times more evidence favouring the null model, suggesting that an underlying pattern does not benefit overall memory performance. This result is consistent with Experiment 2 and Berens et al. (2020), but contrary to Experiment 1.

2.5.2.2. Generalisation

As in Experiment 1, participants showed an ability to generalise, with the distribution of novel clustered locations being significantly less divergent from the von Mises distribution than non-clustered novel locations, $t(124) = 5.00$, $p < .001$, $d = 0.89$, $BF_{01} = 4.66 \times 10^{-5}$ (Figure 2.3C). These results replicate Experiments 1 and 2, showing participants can generalise from old to novel words in the same semantic category.

2.5.2.3. Avoidance

The next analysis tested whether the avoidance effect observed in Experiment 1 would replicate. As shown in Figure 2.3G, participants do show evidence of avoidance behaviour in their location selection. This was confirmed by a significant reduction in the probability density for non-clustered items at the centre of the cluster, $t(62) = 3.35$, $p = .001$ (one-tailed), $d = 0.42$, $BF_{01} = 0.03$ (Figure 2.4C). This replicates the avoidance effect found in Experiment 1. Specifically, participants actively avoid placing the locations of novel non-clustered words at the centre of the cluster.

2.5.3. Discussion

The aim of Experiment 3 was to replicate the findings of Experiment 1, particularly the evidence of an avoidance effect. We found there was no significant benefit to overall mnemonic information available in the clustered compared to the non-clustered condition (as in Experiment 2, but not 1). Second, we replicated the generalisation behaviour seen in Experiment 1. The distribution of locations for novel clustered words was more similar to the underlying von Mises distribution than for novel non-clustered words (Hypothesis 2). Finally, we replicated the exploratory analysis of Experiment 1, showing that participants were less likely to position novel non-clustered words in the centre of the cluster (Hypothesis 3). Experiment 4 aimed to replicate the lack of avoidance following a delay period, as in Experiment 2.

2.6. Experiment 4

Experiment 4 was identical to Experiment 3, apart from the inclusion of a 24-hour delay between Study and Test (as in Experiment 2). We had the same three hypotheses as in Experiment 3. The preregistration for Experiment 4 is available here: https://osf.io/fjze8/.

2.6.1. Methods

2.6.1.1. Participants

2.6.1.1.1. Power Analysis

As before, the required sample size was determined based on the smallest effect size of interest. The minimum effect size of interest (taken across all previous experiments) was d = 0.43 for the total information effect. G*Power (3.1.9, Faul et al., 2007) was used to estimate the required sample size for a paired-samples t-test. Inputting the minimum effect size of interest, α = .05 (two-tailed) and power = .80, suggested a sample size of 45 usable datasets was required. However, a final sample size of 60 usable datasets was set to ensure similar power to previous experiments.

2.6.1.1.2. Final Sample

A total of 79 participants (32 female) with a mean age of 24.15 years (*SD* = 4.54 years) were recruited for the study. Of those, three participants failed attentional checks during Study, three failed to return for the Test phase, five did not provide enough responses, and eight datasets failed to converge during mixture modelling. The final sample was 60 participants (25 female) with a mean age of 24.07 years (*SD* = 4.43 years). All participants were fluent English-speakers with normal or corrected-to-normal vision, were recruited through Prolific.co, and received monetary compensation for their time.

2.6.1.2. Materials and Procedure

The experiment was identical to Experiment 3, except for two features. First, a delay between Study and Test was introduced, similar to Experiment 2. Participants completed the Study Phase and then 24-hours later completed the Test Phase. The average delay was 23.74 hours (*SD* = 0.25 hours). Additionally, participants watched two separate instruction videos, one at the beginning of the Study Phase and another at the beginning of the Test Phase. Written instructions were also provided (https://osf.io/bxru4/).

2.6.1.3. Data Handling and Statistical Analysis

Data handling, exclusion, and statistical analyses were identical to Experiment 3.

2.6.2. Results

2.6.2.1. Memory

There was no difference between the clustered and non-clustered conditions in terms of total information, $t(118) = 1.04$, $p = .299$, $d = 0.19$, $BF_{01} = 3.15$ (Figure 2.2D). As in Experiments 2 and 3, support for the null hypothesis was found.

2.6.2.2. Generalisation

Figure 2.3H shows the pattern of locations selected by participants for novel items in this experiment. It was found that clustered items were significantly less divergent from the von Mises distribution than the non-clustered items, $t(118) = 3.85$, $p < .001$, $d = 0.70$, $BF_{01} = 0.01$. These results replicate all previous experiments.

2.6.2.3. Avoidance

Participants' non-clustered kernel density estimates at the centre of the cluster were compared to a uniform distribution. We found significant evidence of an avoidance effect, $t(59) = 2.00$, $p = .025$ (one-tailed), $d = 0.26$, $BF_{01} = 1.07$. This replicates the findings of Experiments 1 and 3, but not Experiment 2 (where no avoidance effect was present following a 24-hour delay). We return to the possible effect of delay on this avoidance effect in the across-experiment exploratory analyses below.

2.6.3. Discussion

Experiment 4 replicated previous experiments. We found no evidence for a difference in total information between the clustered and non-clustered conditions (as seen in Experiments 2 and 3, but not 1). We showed that the distribution of novel clustered words was more similar to the underlying distribution than for novel non-clustered words (as in Experiments 1-3). We also found evidence that participants were less likely to place novel words in the non-clustered condition near the centre of the cluster (as in Experiments 1 and 3). This was contrary to predictions given the finding of Experiment 2, which found the avoidance effect was no longer apparent following a delay period. To assess this further, we performed an exploratory analysis of the change in avoidance behaviour as a function of time.

2.7. Across-Experiment Exploratory Analyses: All Experiments

Across four experiments, we provide evidence for (1) no difference in overall memory performance (total information) for old words in the clustered relative to the non-clustered condition, (2) less divergence between the experimentally imposed pattern (von Mises distribution) and the novel word responses in the clustered condition relative to the non-clustered condition, and (3) avoidance of the centre of the clustered pattern for non-clustered novel words. We next carried out a set of across-experiment analyses to compare these effects across (1) delay and (2) setting, to ensure the effects are robust to these changes. Further, the individual memory metrics that made up total information (i.e., accessibility and precision) were examined.

Each analysis used a similar GLME structure, assessing whether the metric of interest was affected by Clustering (0 = Non-Clustered, 1 = Clustered), Delay (0 = Immediate Test or 1 = Delay Test) or Setting (0 = In-lab, 1 = Online) along with their interactions. All models, unless otherwise specified, had two random effects per participant. The first was random intercepts per subject, and the other was random slopes for the effect of clustering (if the effect of clustering was included). When no effect of clustering was included, no random slopes were present in the model. Appendix A shows the contrast matrices used for computing the main effects and interactions for these models.

For old words (memory trials), we assessed (1) total information, (2) accessibility and (3) precision, for the clustered condition compared to the non-clustered condition. For new words (generalisation trials), we assessed (1) D_{KL} (relative to the experimentally imposed von Mises distribution) for clustered relative to non-clustered new words, and (2) probability density estimates at the centre of the von Mises distribution for non-clustered new words. When assessing the avoidance effect, no Clustering fixed effect was present in the model.

2.7.1. Memory

2.7.1.1. Total Information

For total information, there was a main effect of delay, $F(1,514) = 48.63$, $p < .001$, $d = 0.39$, $BF_{01} = 5.66 \times 10^{-10}$, with total information decreasing across time. All other main effects and interactions were non-significant ($p \geq .114$, $d \leq 0.12$, $BF_{01} \geq 3.50$). To disentangle this effect, we conducted exploratory analyses of accessibility and precision, the two metrics that make up total information.

2.7.1.2. Accessibility

Figure 2.5, below, shows the effect of clustering, delay and setting on accessibility and precision. For accessibility, there was a significant main effect of clustering, $F(1,514) = 14.24$, $p < .001$, $d = 0.16$, $BF_{01} = 0.02$. Participants showed greater accessibility in the clustered relative to the non-clustered condition. There was also a main effect of delay ($F(1,514) = 32.50$, $p < .001$, $d = 0.39$, $BF_{01} = 1.63 \times 10^{-6}$), with accessibility decreasing from immediate to delayed test. However, no other significant effects were observed ($p \geq .217$, $d \leq 0.07$, $BF_{01} \geq 6.54$).

2.7.1.3. Precision

For the effect of precision, Figure 2.5 shows evidence of clustering having a differential effect compared to accessibility. Specifically, precision appears to be greater in the non-clustered compared to the clustered condition, though only when looking at the online setting. This was corroborated through the analyses performed. There was a main effect of clustering, $F(1,514) = 6.76$, $p = .009$, $d = 0.11$, $BF_{01} = 0.73$. Similar to both total information and accessibility, there was also a main effect of delay, $F(1,514) = 34.20$, $p < .001$, $d = 0.27$, $BF_{01} = 8.28 \times 10^{-7}$, with precision decreasing from immediate to delayed test. A significant interaction between clustering and setting was observed, $F(1,514) = 7.35$, $p = .007$, $d = 0.20$, $BF_{01} = 0.33$. Post-hoc comparisons revealed that the clustering effect was more apparent when the study was conducted online compared to in-lab. To control for familywise error, the Bonferroni-Holm correction was applied and reported as p_{BH}. There was significantly greater precision in the non-clustered condition when tested online compared to the clustered conditions both in-lab (p_{BH} = .031, d = 0.13, BF_{01} = 0.52) and online (p_{BH} = .002, d = 0.17, BF_{01} = 0.03). Additionally, the non-clustered condition showed significantly greater precision when performed online than in-lab (p_{BH} = .024, d = 0.18, BF_{01} = 0.28). All other post-hoc tests were not significant ($p \geq .362$, $d \leq 0.04$, $BF_{01} \geq 12.32$) and no other main effects or interactions were significant ($p \geq .217$, $d \leq 0.07$, $BF_{01} \geq 6.54$).

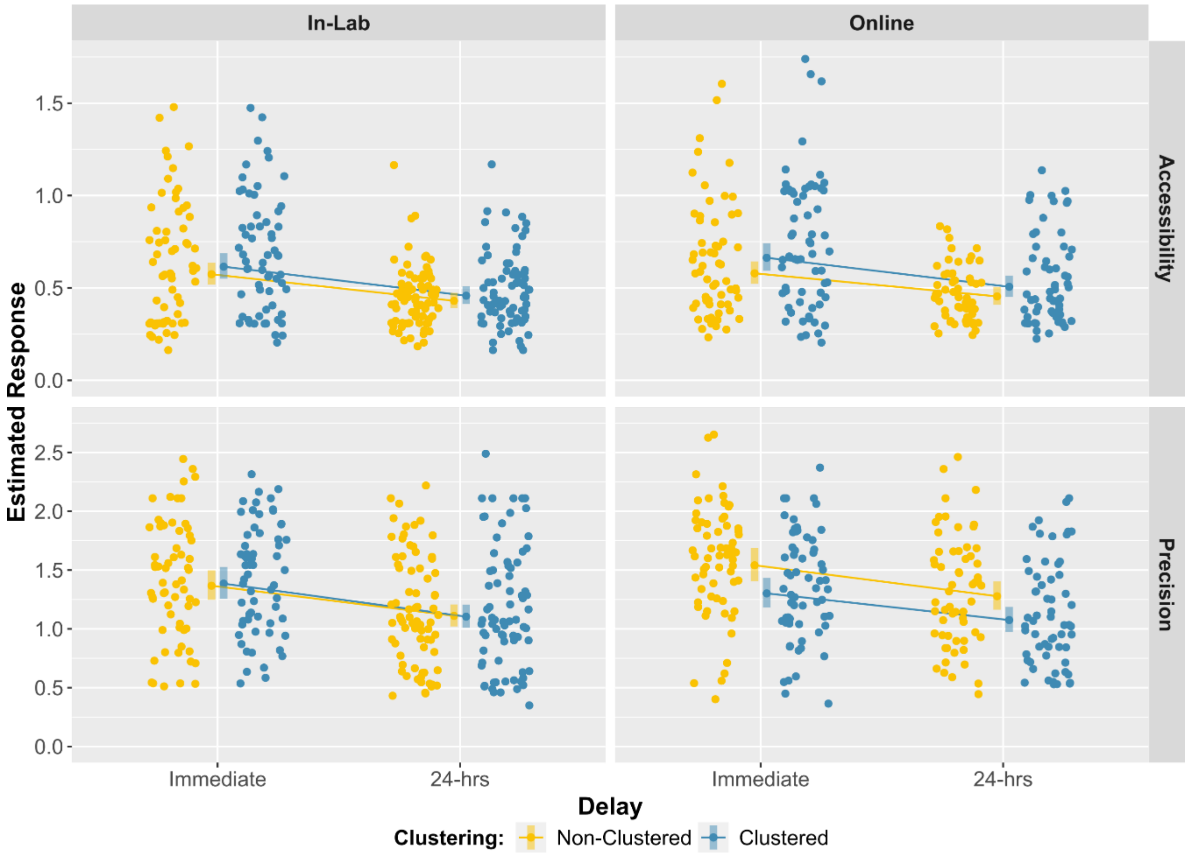

Figure 2.5. **Assessment of accessibility and precision memory measures**. (A) Mean Accessibility (I_p) as a function of clustering, delay and setting. (B) Mean Precision (I_k) as a function of clustering, delay and setting. Individual data points represent participant scores. Error bars represent 95% confidence intervals around the mean for all plots.

2.7.2. Generalisation

2.7.2.1. D_{KL} von Mises

Next, we wished to assess whether the effects of clustering, delay or setting influenced D_{KL} values for the distribution of locations for new words relative to the experimentally imposed von Mises distribution. For this model, we found that only clustering was significant, $F(1,514) = 65.20$, $p < .001$, $d = 0.34$, $BF_{01} = 1.80 \times 10^{-13}$. Specifically, the distribution of clustered new words was less divergent from the von Mises than the distribution for non-clustered new words. All other main effects and interactions were non-significant ($p \geq .051$, $d \leq 0.09$, $BF_{01} \geq 2.88$). Generalisation behaviour was therefore consistent over delay and setting.

2.7.2.2. Kernel Density

For the non-clustered generalisation trials, there was significant evidence of avoidance with reduced kernel density at the cluster centre, $F(1,257) = 28.94$, $p < .001$, $d = 0.17$, $BF_{01} = 1.53 \times 10^{-5}$. No main effects of delay, setting, or an interaction were observed ($p \geq .066$, $d \leq 0.17$, $BF_{01} \geq 1.87$). Therefore, the avoidance effect was consistent across delay and setting.

2.8. General Discussion

Using the precision paradigm, we assessed whether participants use patterns (schematic information) to guide memory and generalisation behaviour. Across four experiments, we found that schematic information modulated both memory and generalisation behaviour. Critically, we found schematic information in one condition (the clustered condition) modulated generalisation behaviour for an unrelated condition (the non-clustered condition). Participants were less likely to place new (generalisation) words in the non-clustered condition near the centre of the clustered pattern. Using the generalisation trials as a proxy, it was possible to assess how schema affect behaviour without the additional noise present for memory trials. Specifically, memory trials can rely on memory for individual items or generalisation based on the underlying pattern. In contrast, generalisation trials are isolated to being influenced by the schematic pattern. Across both exploratory and confirmatory analyses, we found consistent evidence that schematic information influences behaviour for schema-irrelevant information.

2.8.1. Schema-irrelevant Information and Avoidance

The presence of a pattern influenced generalisation behaviour in the non-clustered condition. Participants avoided placing non-clustered items at the location of the cluster for generalisation trials. The presence of schematic information therefore biases generalisation behaviour for schema-irrelevant information.

Previous work has shown that schemas can negatively bias recall of information (Lew & Howe, 2017; Roediger & McDermott, 1995; Warren et al., 2014). For example, Bartlett (1932) demonstrated that retrieval for events in a narrative were biased by a participant's existing knowledge of the world. Warren et al. (2014) showed that, while healthy controls display relatively high levels of false recall in the presence of a schema, patients with vmPFC damage show relatively fewer errors. The Deese-Roediger-McDermott (DRM) false memory effect can also be interpreted as a memory bias (increase in false alarms) in the presence of a schema (Cann et al., 2011). These studies have predominantly focussed on binary measures, demonstrating increased false alarms or errors in the presence of a schema. Here our focus was on information irrelevant to the schematic information being learnt, rather than false memory or biases for schema-related information.

In relation to precision studies, results such as those from Tompary et al. (2020) may have masked this avoidance effect. Tompary et al. (2020) used two clustered conditions on opposite sides of a circle (180^{o} apart), meaning the effects on schema-relevant information will have overshadowed any effects of schema-irrelevant information. It was only with the inclusion of a non-clustered condition, where word-locations were drawn from a uniform distribution that we revealed an effect of the clustered pattern on the semantically distinct non-clustered words.

What produces the avoidance behaviour we observed? One possibility is that the avoidance effect is driven by a “mutual exclusivity” bias (Clark, 1988; Golinkoff et al., 1992). This bias is often studied in language learning and refers to the tendency to only assign one label to an object. For example, suppose children are presented with two objects, one familiar and one novel, and asked to identify what object is being referred to when a novel word is presented. In that case, they typically select the novel object (Markman & Wachtel, 1988); this suggests a reluctance to assign more than one label to a given object, even though several labels may encompass the same object (e.g., a cat is both a mammal and an animal). Though much of the work on mutual exclusivity has focused on children, recent work examining adult word learning has also suggested that the bias helps with generalising to novel words (Lake et al., 2019).

Though this bias is often thought to help guide language development, a similar bias could drive our avoidance effect in the present experiments. As participants identified that semantically related words (e.g., natural words) were associated with a general location (e.g., top-right quadrant), they might have been more inclined to group words from the other category (e.g., human-made) on the opposite side of the circle when generalising. In short, they attributed the top-right quadrant of the circle as “natural only”, despite human-made words also appearing in this area. This explanation is discussed in more detail in Chapter 5.

Another possible explanation for the avoidance effect is a base-rate neglect effect (Hawkins et al., 2015; Welsh & Navarro, 2012; Wolfe, 2007). Specifically, in the non-clustered condition participants’ behaviour may be guided by relative probabilities representing the likelihood of having studied a particular type of word at each location. This contrasts with making location responses based on absolute probabilities representing the overall 'density' of different types of words at each location. Non-clustered word locations were drawn from a uniform distribution, such that the absolute probability of encountering a non-clustered word was close to uniform around the circle. However, the relative probability of encountering a non-clustered relative to clustered word differed around the circle – the relative probability was lower in the clustered area of the circle relative to the other side of the circle. If participants' location responses were influenced by assessing the *relative* probability of having studied a word-location association from a given semantic category, we would expect to observe an avoidance effect in the non-clustered condition. Therefore, when generalising, participants would place non-clustered locations on the opposite side of the circle as non-clustered items are more likely to appear there relative to clustered words.

A third possibility is that the avoidance effect is driven by proactive or retroactive interference between word-location associations (Anderson & Neely, 1996; Baddeley & Hitch, 1977; Barnes & Underwood, 1959; Jenkins & Dallenbach, 1924; Kliegl et al., 2015; Sadeh et al., 2016; Underwood, 1957; Wixted, 2004). Specifically, dense clustering of word-location associations in one part of the circle may result in interference for those specific associations. This interference would apply irrespective of semantic category, resulting in better memory for word-location associations on the opposite side of the circle from the cluster. As these experiences may be more easily retrieved, an apparent avoidance effect for non-clustered generalisation is observed. Specifically, participants during generalisation used more accessible experiences (i.e., those on the opposite side of the circle for non-clustered items), resulting in generalised trials showing an avoidance effect. Under this proposal, the clustered condition would also experience interference. However, the greater number of words located in that area would lead to participants still being more likely to place locations in the clustered area of the circle. This would have the effect of masking the interference in the clustered condition, such that an “avoidance” effect is not seen.

Notably, all three alternative proposals above predict that the avoidance behaviour observed in generalisation should also be present in memory. The mutual exclusivity bias would expect both memory and generalisation trials to be treated similarly, with a lack of assigning non-clustered items to the clustered area. Therefore, non-clustered avoidance should be observed in both conditions. For the base-rate neglect proposal, when an item is forgotten, it assumes that generalisation will occur based on the relative probability of items appearing in a given area of the circle. Therefore, we should observe avoidance in memory, assuming enough items are forgotten. Finally, avoidance in memory should be observed for the interference mechanism because word-location associations within the clustered area are less likely to be retrieved. Consequently, when recalling a location, there may be a tendency to report the item further from the clustered area for non-clustered trials. A formal assessment of memory-based avoidance is presented in more detail in Chapter 3.

2.8.2. Generalisation

Across experiments, participants could use the underlying pattern to make informed decisions about where to locate novel semantically-related words. Distributions across new words in the clustered condition were more similar to the underlying pattern (von Mises distribution) than new words in the non-clustered condition (as measured by D_{KL}).

The evidence for generalisation presented here, both immediately and following a delay, is in line with previous research (Berens & Bird, 2021; Djonlagic et al., 2009; Durrant et al., 2011; Ellenbogen et al., 2007; Graves et al., 2020; Mirković et al., 2019; Sweegers & Talamini, 2014; Tompary et al., 2020). There is conflicting evidence concerning whether generalisation performance increases, decreases, or remains constant over longer delays (Sweegers & Talamini, 2014; Tompary et al., 2020). We saw no clear evidence for a change in generalisation behaviour over a 24-hour delay, suggesting relative stability over one day (which included one night of sleep). Longer delays using a similar experimental approach would be needed to draw definitive conclusions about generalisation behaviour over extended timescales.

The finding of immediate generalisation performance, if such behaviour is based on a schematic representation, is at odds with standard models of systems consolidation (e.g., McClelland et al., 1995). Here, new schematic representations are thought to be formed as a function of hippocampal to neocortical transfer over (at a minimum) several hours, and sleep is thought to play a crucial role in this systems consolidation process (see Rasch & Born, 2013). Although novel information can be rapidly integrated into an existing schema (Fernández & Morris, 2018; Kumaran et al., 2016; van Buuren et al., 2014), this rapid transfer is not thought to occur when establishing new schemas as is the case here, where no location-based schema for a semantic grouping of words should exist before the experiment.

Updated models that incorporate a retrieval-based generalisation mechanism, such as the REMERGE model (Kumaran & McClelland, 2012), more readily accommodate our findings of immediate generalisation. During immediate generalisation, where systems consolidation would not have had chance to take place, participants will rely more on retrieval-based generalisation mechanisms. Over time, as systems consolidation occurs, there will be a move to more encoding-based mechanisms supported by a generalised neocortical-based schema.

It is plausible that there is a shift from retrieval-based to encoding-based generalisation over time in our experiments, but that both mechanisms support similar generalisation behaviour. However, recent research suggests generalisation behaviour might decrease over time, which would be inconsistent with the extraction of a stable schematic representation. Using a similar paradigm, Tompary et al. (2020) showed that schematic representations may decline over time alongside memory for individual word-location associations. Antony et al. (2021) found a similar pattern of results using a spatial navigation object-location task. As participants' memory performance declined for individual object-location associations over time, so did their adherence to the pattern of locations. Finally, although they did not assess generalisation to new words, Berens et al. (2020) showed that the distribution of remembered word locations decreased in similarity to the underlying pattern over four days and this decrease correlated with memory accessibility (i.e., the proportion of word-location associations retrieved). These results are more in line with retrieval-based generalisation and may suggest the generalisation observed in the present work uses this same approach.

2.8.3. Memory

The schematic information in the clustered condition modulated memory-guided behaviour in both the clustered and non-clustered conditions. First, in an exploratory analysis, we replicated the results of Berens et al. (2020), showing the presence of a pattern increased accessibility (proportion remembered) but decreased precision (the angle of error for word-location associations that were remembered). Our pre-registered analyses comparing Total Information (the product of accessibility and precision, divided by a constant) in the clustered relative to non-clustered condition showed no overall boost in memory performance between conditions (though a small but significant difference was seen in Experiment 1). This lack of an increase in overall memory performance again replicates the results of Berens et al. (2020).

Previous studies have shown an overall benefit to memory for schematic vs non-schematic information (Atienza et al., 2011; Brewer & Treyens, 1981; Frank et al., 2018; Greve et al., 2019). The present findings might appear to contradict these studies. However, most previous analyses have used binary measures of memory (correct vs incorrect) that are conceptually similar to the accessibility measure used in the present studies. Thus, our increase in accessibility in the clustered relative to non-clustered condition is consistent with previous findings.

Importantly, our ability to assess accessibility and precision suggests this increase in accessibility comes at a cost – a corresponding decrease in precision. This reduced precision is similar to previous findings suggesting that the presence of a schema leads to the loss of more fine-grained detail information, but enhanced memory for face-location associations that had a schematic element (Sweegers et al., 2015). Other studies have reported similar memory biases as a consequence of schematic information (Berens et al., 2020; Mäntylä & Bäckman, 1992; Pezdek et al., 1989; Richter et al., 2019; Tompary et al., 2020; Tompary & Thompson-Schill, 2021). Therefore, our results are consistent with previous studies that schematic information can increase performance on certain measures of memory, but decrease performance on others.

Further, our findings concerning accessibility and precision suggest that the increase in "information" in terms of accessibility is equivalent to the decrease in precision (hence the lack of difference in Total Information), such that schematic information in this paradigm does not increase overall memory performance. Although we cannot yet generalise beyond the present experimental approach, one possibility is that this accessibility versus precision trade-off (or the trade-off between hits and false-alarms in other experiments) might result in no net memory benefit in the presence of a schema. In short, schematic information alters memory behaviour, but our results question whether they benefit overall memory performance.

2.8.4. Conclusion

Across four experiments, we provide evidence for memory and generalisation effects for both schema-relevant and -irrelevant information. Critically, we have shown that generalisation behaviour is biased away from a schematic location for schema-irrelevant information. These effects appear immediately after encoding and appear to be relatively stable over a 24-hour period. We have outlined three broad explanations for this behaviour outside of schema-based processes: (1) a mutual exclusivity bias account, (2) a base-rate neglect account and (3) an interference account.

Given these effects emerge immediately after encoding, with evidence of decline over longer delays in other experiments (e.g., Antony et al., 2021; Tompary et al., 2020), the generalisation behaviour is likely driven by a retrieval-based mechanism that infers a location based on a “on the fly” retrieval of word-location associations that are semantic neighbours to the novel item. Formal modelling is likely to provide further theoretical insight (see Chapter 3). For example, accessibility and precision measures have recently been suggested to emerge from a single *d*-prime measure in a signal detection framework (Schurgin et al., 2020). Incorporating both location-based interference and semantic relatedness in such a framework may be able to accommodate our findings without the need for schematic representations or semantic categorisation.

Chapter 3:
Exploring the Avoidance Effect: Using computational modelling to investigate the avoidance effect

A secondary data analysis using the data collected in Berens et al. (2020) is reported in this chapter. The data for this analysis can be found on the Open Science Framework: https://osf.io/6mx3s/. The pre-registration, data and analyses scripts used for the analyses reported can also be found on the Open Science Framework: https://osf.io/bxru4/.

The exploratory and confirmatory analyses reported in this chapter are published as part of a preprint: Cockcroft, J. P., Berens, S., Gaskell, M., & Horner, A. J. (2021, August 24). *Schematic information influences memory and generalisation behaviour for schema-relevant and -irrelevant information*. https://doi.org/10.31234/osf.io/nzurq. The content was moved to the present Chapter to increase continuity across chapters.

3.1. Abstract

The present Chapter aimed to explore why the avoidance effect was present during the behavioural experiments reported in Chapter 2 using computational modelling. The first family of models assessed whether simple encoding- or retrieval-based models could predict the presence of the avoidance effect. However, neither model produced this effect, demonstrating that in their basic form, neither encoding- nor retrieval-based models are sufficient to explain the avoidance behaviour. Subsequently, the second family of models examined how modulating retrieval probability, whereby items in the clustered condition were better remembered within the cluster and non-clustered items were better remembered further away, would predict the avoidance effect. This model did generate an avoidance behaviour during generalisation, similar to the behavioural experiments in Chapter 2. However, it also generated the same avoidance behaviour in memory; this was something that had not been formally explored. Exploration of the data from Chapter 2, and analysis of an independent dataset, supported the model's prediction that avoidance was also present during memory. The final model explored how location- and semantic-based interference influenced behaviour. It was found that this model provided a good fit for the data, providing a more parsimonious solution than the previous models. Consequently, results consistent with the presence of schemas (e.g., Chapter 2) may be explained by a non-schematic model of memory.

Keywords: *encoding-based, retrieval-based, interference, avoidance*

3.2. Introduction

In Chapter 2, we observed an unexpected effect when it came to generalisation. Specifically, participants showed an active avoidance of placing non-clustered novel locations within the clustered area of the circle. This surprising result was found in both exploratory (i.e., Experiment 1) and confirmatory (i.e., Experiments 3 and 4) analyses. The present Chapter aimed to better understand this avoidance effect by examining different mechanisms explaining its presence.

3.2.1. Encoding- vs Retrieval-Based Models

As discussed in the Literature Review (see Chapter 1), generalisation could occur through encoding- or retrieval-based mechanisms (Hintzman, 1986; McClelland et al., 1995; Nosofsky, 1988; Rosch, 1973). Typically, encoding models propose that at the point of encoding, overlapping events create abstract representations (or schema) that are used when generalising. In contrast, retrieval-based models argue that the individual events are sampled to generalise “on the fly”. In this way, the two models differ in how memories of the events are relied upon during generalisation. Encoding models propose that individual experiences are not relied upon but instead a generalised representation (e.g., the average of experiences) is. In contrast, retrieval models propose that the individual experiences are relied upon without needing a schematic representation.

In their basic forms, neither an encoding nor retrieval-based model seemingly predicts the avoidance behaviour observed in Chapter 2. An encoding-based model would predict that the central tendencies across events should be extracted and used to generalise. As such, the clustered condition should remap the von Mises distribution (i.e., the pattern underlying the cluster), whilst the non-clustered should remap the uniform distribution. Though clustered generalised items did remap onto the von Mises distribution, the non-clustered did not remap the uniform. For the retrieval-based model, sampling from experiences should mean the individual also remaps the presented pattern. Therefore, like the encoding-based approach, the retrieval-based model would predict a cluster in the clustered condition and a uniform density in the non-clustered. The first family of models (see Model Family 1, below) confirms this prediction, showing that neither a "simple" encoding- nor retrieval-based model predicts the avoidance effect in the non-clustered condition.

3.2.2. Mechanisms of Avoidance

3.2.2.1. Schema and False Memory

As discussed in Chapter 1, schema presence may distort memory for events (Aizpurua et al., 2009; Bower et al., 1979; Brewer & Treyens, 1981; Nakamura et al., 1985; Yamada & Itsukushima, 2013). Therefore, we may observe the avoidance behaviour in the non-clustered condition as any schema formed within the clustered condition subsequently influenced retrieval for schema-relevant (clustered) and -irrelevant (non-clustered) information. Within the clustered condition, there was the possibility of a schema developing due to the overlapping patterns across events (Ghosh & Gilboa, 2014; van Kesteren et al., 2012). Subsequently, any schema developed for the clustered condition may have biased memory for non-clustered items by reducing the probability of retrieving non-clustered items when they appear within the clustered area. This decrease in probability may relate to these items being incongruent with expectations (i.e., the idea that a given area of the circle is devoted to one set of object nouns). People may then generalise based on memories with a higher retrieval probability (i.e., those congruent with the schematic representation). The consequence of relying on those experiences when generalising is the formation of an avoidance effect for non-clustered items, as sampled experiences for that condition are more likely to appear outside the clustered region. The second family of models (see Model Family

2, below) demonstrates how modulating the retrieval probability of clustered and non-clustered words according to their location around the circle can produce an avoidance effect in the non-clustered condition.

3.2.2.2. Interference

An alternative mechanism is interference (Anderson & Neely, 1996; Baddeley & Hitch, 1977; Barnes & Underwood, 1959; Kliegl et al., 2015; Underwood, 1957). Interference can occur both retroactively and proactively. Retroactive interference occurs when newly encoded items decrease the probability of retrieving older information, whilst proactive interference occurs when existing memories decrease the probability of retrieving newly learned information. In their study, Baddeley and Hitch (1977) had rugby players recall the names of teams they played against throughout the rugby season. It was found that as the number of games increased, the recall of team names decreased; this was irrespective of the time that elapsed. Therefore, as more games were played, the probability of recalling the names of teams previously played against decreased; this is an example of retroactive interference. An example of proactive interference is shown in Underwood (1957). In their study, Underwood (1957) asked participants to learn nonsense syllables. Participants were either in the interference or control condition. In the interference condition, participants had previously learned a list of nonsense syllables. For the control condition, no previous learning had taken place. Those in the interference condition showed only 20% recall of the new list items after a 24-hour delay, compared to 80% recall for the control condition. Consequently, the learning of the first list reduced the probability of retrieving the new list of syllables.

There is an ongoing debate about whether interference occurs at the point of encoding or retrieval (see Kliegl et al., 2015). Some authors attribute interference to be a process at encoding, whereby as attention decreases and memory load increases, interference takes place (e.g., Crowder, 1976). In this way, earlier memories will be better remembered due to better encoding, whilst later memories will not. In contrast, retrieval accounts, such as temporal discrimination theory (Wixted & Rohrer, 1993), propose that interference results from being unable to disentangle contexts across learning, leading to memory searches that retrieve nontarget experiences. More recent accounts and evidence have tried to support the view of a hybrid encoding-and-retrieval based account of interference (Kliegl et al., 2015). Kliegl et al. (2015) argue that interference builds at both encoding and retrieval. At encoding, both attentional disengagement and heightened memory load result in encoding-based interference. Subsequently, at retrieval, during memory search, both target and other related information will be retrieved, leading to further interference.

One account of interference that may explain the avoidance effect observed in Chapter 2 is the cue-overload principle (Watkins & Watkins, 1975). The cue-overload principle refers to the idea that retrieval probability decreases as the number of items associated with a cue increase. According to this principle, competing information associated with the target area will be activated during a memory search resulting in interference. This account acts similarly to notions of spreading activation (e.g., Anderson & Pirolli, 1984). Applying this to the avoidance effect, the probability of retrieving an item within the clustered area may decrease the more items that appear there, irrespective of condition. Therefore, when retrieving locations from memory, an individual may be more able to access memories away from the clustered area due to the reduced competition. Under this assumption, there may be an avoidance in the clustered condition. However, given the sheer number of clustered items that appear there, in contrast to the non-clustered items, an avoidance is not observed. However, for the non-clustered condition, retrieval probabilities will be much lower in the circle's clustered area than the other side of the circle, leading to an apparent avoidance effect. Of course, this proposal would also predict avoidance in memory (this is assessed later in the chapter).

One of the benefits of the interference-based proposal is that it can occur either explicitly or implicitly. During the debrief of Chapter 2, very few (10/261, 3.83%) participants explicitly mentioned using a strategy where one group of words were placed in one area and the other group on the opposite side of the circle. Further, few (4/261, 1.53%) participants explicitly labelled items at the superordinate level (i.e., human-made vs natural). Instead, participants focused on the basic categories (e.g., animals, office furniture). As such, given that the interference mechanism would modulate retrieval probability irrespective of condition, there is no need for an explicit strategy of: “If natural, place here. If not, place on the opposite side”. The final model (see Model 3, below) demonstrates how a location- and semantic-based interference mechanism can explain the avoidance effect in the non-clustered condition.

3.3. Overview of Models

This Chapter aimed to assess how different mechanisms may explain the avoidance effect. Across models, different approaches to derive location judgements for both “memory” and “generalisation” trials were applied. These models generated responses by acting as “participants” encoding and retrieving locations when completing the precision task. Throughout the chapter, the generated responses from the model are referred to as “participants”. Each model generated responses in the task using a different mechanism. The first family of models investigated the locations selected using basic forms of encoding and retrieval-based generalisation models as means of assessing whether they predicted an avoidance behaviour during generalisation. The second family of models implemented a retrieval probability mechanism, where clustered items had a higher probability of being remembered closer to the cluster, whilst non-clustered items were less likely to be remembered in that area. Finally, the third model used both location- and semantic-based interference to affect location judgements.

All models presented throughout this Chapter were developed in MATLAB (2019b) using in-built functions along with the HoopStats toolbox (v. 2.00; Berens et al., 2020) and Circular Statistics toolbox (v. 1.21; Berens, 2009). All models generated data for 1,000 simulated "participants".

3.3.1. Model Input and Output

All models received target locations to use as a basis for generating "remembered" and "generalised" locations. Target locations were generated in the same way as the behavioural experiments described in Chapter 2 by creating a random sample of locations based on von Mises parameters (μ and κ). The von Mises distribution is a continuous probability distribution within a circular space analogous to the Gaussian distribution (Best & Fisher, 1979). The probability density function is shown below (1), with I_0 representing the modified Bessel function of order 0. The μ parameter represents the mean location and κ the distribution's concentration (i.e., width). As with the behavioural experiments, the clustered condition used the parameters: μ = 0 and κ = 2, whilst the non-clustered condition had parameters: μ = 0 and κ = 0. Setting the concentration to 0 ensured the sampled locations were from a uniform distribution. A total of 180 target locations were generated, 90 for each condition.

$$f(\theta \mid \mu, \kappa) = \frac{e^{\kappa \cos(\theta - \mu)}}{2\pi I_0(\kappa)} \quad (1)$$

The output from the models were locations selected for both memory and generalisation trials. Across models, items were identified as remembered or forgotten, though, this was done differently across models. For example, the second family of models determined remembered and non-remembered items based on a retrieval probability. Further, the location chosen when an item was “remembered” varied based on the model. For instance, in the interference model (i.e., Model 3), item locations were selected by weighting a location judgement by the probability of retrieving that item and its semantic similarity to other items. In this way, the remembered location would not be an exact replacement of the target but based on “noise” created from other associated items and its retrieval probability. Across models, generalised locations were generated using either an encoding- or retrieval-based approach. For instance, taking the remembered items and fitting a von Mises distribution (i.e., a “schema” of events) or selecting a sample of remembered items and taking an average.

3.3.2. Model Assessment

Model assessment used similar measures from the behavioural data. Specifically, for the memory trials, I_p and I_k parameters were computed to assess memory accessibility and precision. The values are expected to replicate the pattern of findings from Chapter 2, meaning accessibility should be greater in the clustered compared to the non-clustered condition. However, precision should be greater in the non-clustered compared to the clustered condition. The mixture model procedure described in Chapter 2 was applied to the model responses to generate the I_p and I_k parameters. Briefly, the mixture model applied an expectation maximisation algorithm to estimate retrieval probability (p) and precision (k) from the angular error (i.e., the difference from the target location to the selected location). These measures were then converted into entropy measures of I_p and I_k.

Kernel density estimates were used to assess for evidence of generalisation (in the clustered condition) and avoidance (in the non-clustered condition). These were estimated similarly to Chapter 2. Briefly, the kernel density estimates represent a smoothed probability density function for the locations selected by the model. A von Mises density function, with parameters $\mu = 0$ and $\kappa = 2$, acted as a smoothing kernel for each response around the circle. These densities provide estimates of the average probability that "participants" selected a specific location around the circle. The kernel density values are expected to differ from the uniform density value ($2\pi^{-1} = 0.159$). For clustered items, the values should be greater than the uniform, whilst the non-clustered items will either be close to (if locations are uniformly distributed) or less than (if an avoidance effect is present) the uniform. Along with being used for model assessment, these density estimates were used to plot the average locations selected by the model for both memory and generalisation trials, similar to Chapter 2.

It is worth noting that no statistical models were generated to examine the effects of interest. Given that it was possible to control the amount of variance present in the generated data by changing the number of trials or "participants", applying statistical models was deemed inappropriate. In other words, if an "effect" is present, it can trivially be made "significant" by increasing the number of iterations. Therefore, when producing results for the models, only the mean values are reported. However, examining whether the mean values produced by the model mapped onto the data from Chapter 2 was undertaken. The models did not necessarily aim to remap the mean values from Chapter 2, but instead adhere to the general pattern of behaviour (e.g., increased accessibility in the clustered compared to non-clustered condition). The descriptive statistics for Chapter 2 data are shown in Table 3.1 below. Note that these are not values derived from the GLME's reported in Chapter 2 but from the data itself.

Table 3.1.

Descriptive statistics for Chapter 2 isolated to the dependent variables of interest.

Variable	Group	*M*	*SD*	*SE*	95% CI Lower	95% CI Upper	Diff
Accessibility (I_p)	NC	0.542	0.262	0.016	0.511	0.573	-0.061
	CL	0.603	0.292	0.017	0.569	0.637	
Precision (I_k)	NC	1.365	0.499	0.030	1.306	1.423	0.098
	CL	1.267	0.486	0.029	1.210	1.324	
Kernel Density: Memory	NC	0.155	0.040	0.002	0.151	0.160	NA
Kernel Density: Generalisation	NC	0.148	0.055	0.003	0.142	0.155	NA
	CL	0.184	0.060	0.004	0.177	0.191	

Note. 95% Confidence Intervals were computed using the Student's *t*-Distribution. NC = Non-Clustered, CL = Clustered. Diff = Difference between mean values, subtracting CL from NC. NA = Not applicable.

3.3.3. Controlling Behavioural Performance

Behaviour was modulated in the models based on the behaviour observed in Chapter 2 using the p and k variables (as opposed to I_p and I_k). Beta and gamma parameter estimates were generated for the p and k variables using maximum likelihood estimation. The use of p and k occurred as this provided us estimates of proportion remembered (p) and levels of precision (k), which could be used when generating responses. Though I_p and I_k are the same metrics on a different scale, they cannot be easily used within existing functions to generate proportion remembered and concentration without converting them back to their original metric space.

A beta (p) and gamma (k) distribution were fit to each measure separately for clustered and non-clustered conditions (see Table 3.2, below) using all 261 participants data described in the previous chapter. As p was bounded between 0 and 1, a beta distribution was the most appropriate, whilst k could take on values from 0 to positive infinity, meaning a gamma distribution was the best fit. These values were used in the first two models to control the performance of "participants" in the task.

Table 3.2.

Maximum likelihood estimates for both the p and k parameters.

Variable		Distribution			
		Beta		Gamma	
		α	β	k	ϑ
p	NC	3.09 [2.56, 3.72]	7.11 [6.23, 8.12]	-	-
	CL	2.66 [2.23, 3.16]	5.21 [4.59, 5.91]	-	-
k	NC	-	-	1.37 [1.18, 1.60]	8.22 [6.82, 9.01]
	CL	-	-	1.45 [1.24, 1.70]	6.45 [5.36, 7.76]

Note. Values included in square brackets represent the 95% CI of the parameter estimate.
NC = Non-Clustered, CL = Clustered.

3.4. Model Family 1: Basic Encoding vs Retrieval-Based Generalisation

3.4.1. Methods

3.4.1.1. Memory

The first family of models examined what encoding and retrieval-based models predicted for the locations selected during generalisation. Each "participant" had a random p and k value selected, which would form the basis of that "participant's" performance on the task, ensuring they only remembered a proportion (p) of events at a certain degree of precision (k). The p values sampled did not go below 8, as per the requirements of the mixture model algorithm to ensure the mixture model would fit the output data (see Chapter 2).

Using the selected k parameter, a new set of "remembered" locations were generated via a von Mises distribution with a mean of 0 and the participant's selected level of precision. These randomly generated angles were added to the target values; this provided participants' responses that adhered to their selected level of precision. These responses were categorised as either "remembered" or "non-remembered"; this was done for two reasons: (1) the non-remembered values would be changed to random guesses, and (2) for the encoding and retrieval-based models during generalisation. Trials were classified as either remembered or non-remembered so that the proportion of remembered trials was equal to the participant-specific p parameter. For the responses categorised as non-remembered, a set of random angular locations were generated using the von Mises parameters: μ = 0 and κ = 0. These values would then replace the responses categorised as "non-remembered" to act as "random guesses" and have no relation to the previously generated target locations. In this way, each target location was associated with a "remembered" or "non-remembered" response. If it is not remembered, the retrieved location is a random circular location. If it is remembered, the retrieved location is the target location with the addition of "noise".

3.4.1.2. Generalisation

Once all memory responses had been generated, generalised responses were computed adhering to either an encoding or retrieval-based protocol.

3.4.1.2.1. Encoding Model

Encoding models generally predict that, at the point of encoding, regularities across events are extracted to form long-term schematic representations. These representations are then used when making novel inferences. For the present models, a schema was generated by fitting a von Mises distribution to a set of angular responses. Three versions of an encoding model were developed, with the critical difference related to the responses used. For the first model (Figure 3.1, E.1), all memory responses within a given condition were used regardless of whether they were classified as "remembered" or "non-remembered". The second model (Figure 3.1, E.2) used a perfect encoding strategy, whereby all targets for a condition were input to generate the parameters. Finally, the third model (Figure 3.1, E.3) used only the memory responses categorised as remembered. The parameters derived from these responses were then used to generate a random sample of angles based on the fitted μ and κ values; these acted as the generalised responses for a given participant.

3.4.1.2.2. Retrieval Model

In contrast to encoding models, retrieval models generally argue that generalisation relies on a sample of individual memories to generate a response; this occurs “on the fly”. For the present modelling work, three types of retrieval-based models were generated similar to the encoding models described above. However, all models used only trials identified as “remembered”. For all models, three experiences were sampled; this was chosen as changing the number of sampled experiences between 2 and 10 did not appear to affect model output. The first model (Figure 3.1, R.1) used the mean location from the sampled experiences. The second model (Figure 3.1, R.2) selected one of the sampled experiences and used that as the response. Like the encoding model, the final model (Figure 3.1, R.3) estimated the von Mises parameters based on the sampled experiences; this would be a noisy estimate as few responses were retrieved. The generated parameter estimates were then used to make a generalised response.

3.4.2. Results

Figure 3.1 below shows the spatial distribution of locations selected for memory and generalisation trials, centred to the von Mises distribution. As shown, all models produced the output matching their respective distribution. For the clustered items, there was an increased density of locations towards the centre of the cluster. In contrast, for non-clustered items, responses were uniformly distributed across the circle. However, there was no apparent evidence of avoidance in the non-clustered condition.

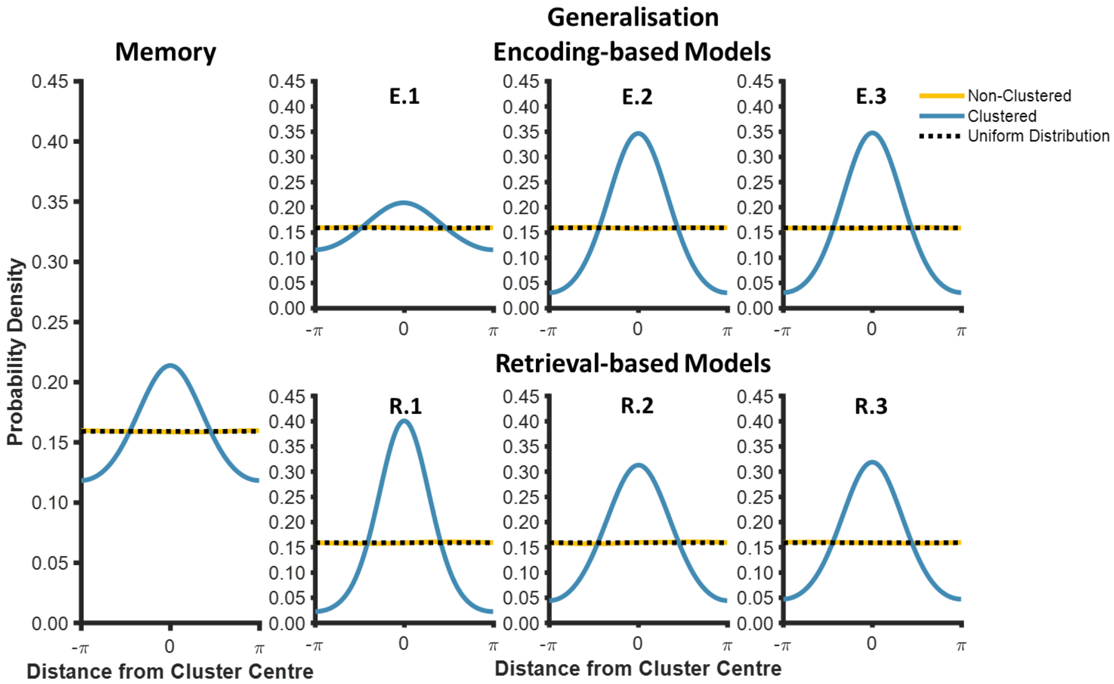

Figure 3.1. **Model Family 1: Density of Locations**. Plots show the distribution of locations selected by the model for both memory and generalisation items, centred to the experimentally imposed von Mises distribution. On the far left are the memory trials. To the right of that are generalisation trials. The top row shows the encoding-based models **E.1: All Locations Model**. **E.2: Target Locations Model**. **E.3: Remembered Locations Model**. The bottom row shows retrieval-based models. **R.1: Mean Location Model**. **R.2: Pick Sampled Experience Model**. **R.3: Fit Distribution Model**. The dashed line on all plots represents the uniform probability density value (0.159).

3.4.2.1. Memory

Table 3.3 below shows the mean estimates for the variables of interest. For accessibility (I_p), the mean estimates from the model follow a similar pattern to the behaviour observed in Chapter 2, whereby the clustered condition had greater levels of accessibility than the non-clustered. Similarly, levels of precision (I_k) followed the pattern of behaviour from Chapter 2, with greater precision in the non-clustered than the clustered condition. However, the difference between conditions for precision was much smaller using the model than was found in the actual data. These models were provided with separate p and k values for the clustered and non-clustered conditions according to the values determined from Chapter 2. Therefore, this difference in accessibility and precision is not a result of the mechanics of the model but the input to it (c.f. Model 3).

Table 3.3.

Descriptives for memory variables from Model Family 1, including whether the value was within the 95% CI of the Chapter 2 data and the differences between conditions.

Variable	Condition	*M*	Within 95% CI?	Group Difference
I_p	NC	0.572	Y	-0.053
	CL	0.637	Y	
I_k	NC	1.420	Y	0.019
	CL	1.401	N	
PKD: Memory	NC	0.159	Y	NA

Note. I_p = Accessibility, I_k = Precision. PKD: Memory = Kernel density at cluster centre for memory trials. NC = Non-Clustered, CL = Clustered. Y = Yes, N = No. Group difference was calculated by subtracting the CL mean from the NC mean. NA = Not applicable.

3.4.2.2. Generalisation

As shown in Figure 3.1 above, there was evidence of the model being able to generalise from “memory” to “generalisation” trials, adhering to the patterns presented at encoding. Specifically, the model correctly placed clustered items within the cluster whilst treating the non-clustered as having an equal likelihood of appearing throughout the circle (i.e., uniformly distributed). Table 3.4, below, provides the mean values from the model. When comparing to the output from Chapter 2, it is clear the model over-predicted the number of responses at the cluster centre for both non-clustered and clustered trials, with all values being outside the 95% CI’s. Further, the non-clustered condition showed essentially no difference from the uniform density (0.159); this suggests the models lacked evidence of an avoidance behaviour.

Table 3.4.

Descriptives of generalisation variables from Model Family 1, including whether the value was within the 95% of the Chapter 2 data.

Model	Condition	*M*	Within 95% CI?
Encoding Model 1: All Responses	NC	0.159	N
	CL	0.209	N
Encoding Model 2: All Targets	NC	0.158	N
	CL	0.347	N
Encoding Model 3: Remembered Only	NC	0.159	N
	CL	0.348	N
Retrieval Model 1: Circular Mean	NC	0.159	N
	CL	0.401	N
Retrieval Model 2: Choose One	NC	0.160	N
	CL	0.313	N
Retrieval Model 3: von Mises Fit	NC	0.159	N
	CL	0.319	N

Note. NC = Non-Clustered, CL = Clustered. Within 95% CI = Whether the value was within the 95% CI of Chapter 2. Y = Yes, N = No.

An analysis was undertaken to assess whether altering the number of samples used for the retrieval-based models resulted in the presentation of avoidance. No evidence of change was identified based on the number of samples, with the mean values being identical to the output described above.

3.4.3. Discussion

This family of models assessed whether variants of the encoding or retrieval-based model would predict the behaviour observed in Chapter 2, particularly the avoidance effect. The model generated memory responses with greater accessibility in the clustered compared to the non-clustered condition. Similarly, there was evidence of greater precision in the non-clustered condition compared to the clustered. However, as these parameters were within the model itself, these results are not too surprising. Nevertheless, the model generated clustered generalised responses that showed greater density at the cluster centre, evidencing adherence to the von Mises distribution. Therefore, both an encoding- and retrieval-based model could explain generalisation performance in the clustered condition. However, there was no evidence of avoidance for non-clustered generalised items, with the mean values being identical or in proximity to the uniform value; this is further supported by examining the distribution of locations selected for generalisation trials (see Figure 3.1).

From this family of models, it was clear neither encoding- nor retrieval-based models alone would predict an avoidance behaviour during generalisation. This was not a surprising result but further corroborates arguments mentioned during the Introduction. Specifically, given either an encoding- or retrieval-based process, the pattern of responses should follow the underlying distribution presented during encoding, with some amount of noise. This noise would not result in an average bias in the locations selected for generalised non-clustered trials, which themselves are based on a sample of uniform angles. Therefore, this set of models provides formal evidence that simple encoding- or retrieval-based approaches cannot predict the avoidance effect for non-clustered generalised responses.

One caveat with the present family of models is how items are categorised as remembered or non-remembered. Currently, the models ignore the target proximity to the underlying cluster. Research has shown that retrieval is modulated by the extent to which items map onto the underlying pattern (e.g., Tompary et al., 2020; Tompary & Thompson-Schill, 2021). For example, in their categorisation study, Tompary and Thompson-Schill (2021) found that recall of locations for items was biased towards the mean location a target item belonged to. Therefore, not accounting for the target locations proximity to the cluster may misrepresent the data.

Further, the model assumes that participants differentiate between old and new items with 100% accuracy. Specifically, generalisation would only be present for generalised trials. In contrast, memory trials regarded as “non-remembered” would be randomly guessed. This seems counterintuitive and goes against previous work demonstrating that recognition for old-new items is not always accurate, particularly when items are semantically related (Deese, 1959; Montefinese et al., 2015; Roediger & McDermott, 1995; Shiffrin et al., 1995). As such, it is possible that when participants have a low probability of remembering an individual item, they rely on the same generalisation processes for memory trials as they do when generalising to novel instances.

3.5. Model Family 2: Modulating Retrieval Probability

The primary aim of this model was to assess how modulating retrieval probability based on proximity to the cluster could result in the presence of an avoidance effect. Specifically, using a retrieval probability function, the probability of retrieving an item was modulated so that clustered items were better remembered closer to the cluster, whilst non-clustered items were better remembered further away. By modulating retrieval probabilities in this way, an avoidance during generalisation may be observed as more accessible non-clustered memories (i.e., those with higher retrieval probabilities) are further from the clustered location. Along with implementing the retrieval probability function, the model used generalisation mechanisms when memory responses were classified as “non-remembered” instead of randomly guessing (contrary to Model Family 1).

3.5.1. Method

The model worked similarly to Model Family 1, with some exceptions as outlined below.

3.5.1.1. Memory

The model determined remembered and non-remembered items based on a retrieval probability function developed by S. C. Berens (personal communication, 20 August 2020). These probabilities were estimated so that items closer to a specified mean value (e.g., 0) were better remembered than those further from this value. For the clustered condition, this would be the mean value of the von Mises distribution. For the non-clustered condition, the retrieval probability mean was set 180° away from the clustered mean (i.e., on the opposite side of the circle). The retrieval probability function used the individual "participant's" p value to determine the number of items that should have higher retrieval probabilities. More specifically, probabilities were estimated such that they were sigmoidal to the probability density function of a von Mises distribution, with parameters μ and κ. The integral of these values was exactly equal to the expected p value. Therefore, if the von Mises distribution had a μ of 0 and κ of 2, then items closer to 0 were given higher retrieval probabilities than those further away, with the individual p value for each "participant" being used to estimate how many items were to be remembered. An infographic of the function is shown below (see Figure 3.2). The formula for this function are provided in Appendix B.

Figure 3.2. **Infographic of the Retrieval Probability Function**. This illustrates how the retrieval probability function would work, whereby filled dots represent items classified as having a greater probability of being remembered, whilst unfilled dots represent a lower probability of being

remembered. Non-clustered items (yellow) were more likely to be remembered further from the cluster. In contrast, clustered items were more likely to be remembered if they appeared within the clustered area.

Once the retrieval probabilities for each trial were determined, items were explicitly classified as "remembered" or "non-remembered" based on whether the assigned probability value differed from a randomly generated number between 0 and 1. If the retrieval probability value was greater than or equal to the random value, the item was categorised as "remembered"; otherwise, it was "non-remembered". For those items classified as "non-remembered", the model did not generate random responses as was done in the first family of models, but instead determined locations via generalisation (i.e., treating them as "novel" items). The generalisation mechanisms are described in more detail below.

3.5.1.2. Generalisation

Unlike the previous model, which used three types of encoding- and retrieval-based generalisation, the present family of models only used one type; this was due to all versions in the previous model producing similar behaviour. For the encoding-based model, a von Mises probability density function was fit to items classified as "remembered" in order to estimate the μ and κ values. These parameters were then used to generate generalised responses. For the retrieval-based model, a sample of remembered responses was taken, with their retrieval probability estimates providing a weighting as to which items would be sampled. Therefore, items with higher retrieval probabilities were more likely to be sampled when a location judgement for novel (or non-remembered) items occurred. A total of three items per participant were sampled. The mean of these sampled locations was then calculated and used as a response.

3.5.2. Results

Figure 3.3 below shows the distribution of locations selected for memory and generalisation trials centred to the experimentally imposed von Mises distribution. As shown, the model appears to have produced an avoidance effect during generalisation for both encoding and retrieval-based models. However, an avoidance behaviour was also present during memory trials; the presence of an avoidance behaviour in memory has not been formally explored in our behavioural data.

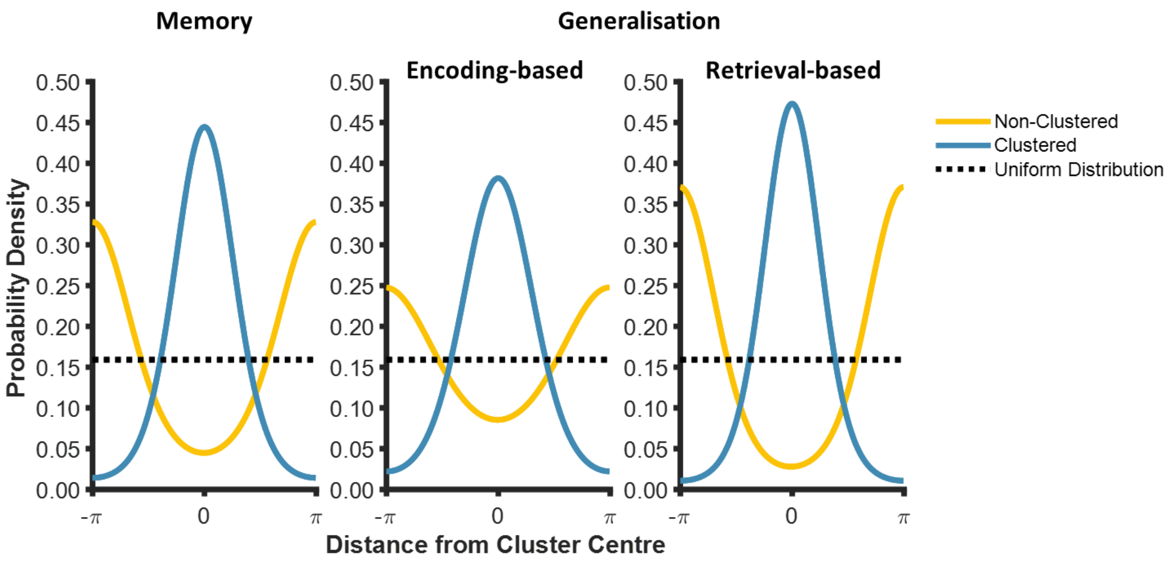

Figure 3.3. **Model Family 2: Density of Locations**. Plots show the distribution of locations selected by the models for memory (left), encoding-based generalisation (middle) and retrieval-based generalisation (right) when centred to the experimentally imposed von Mises distribution. The density of locations towards the cluster centre is greater in the clustered condition, with a decreased density for non-clustered items (i.e., an avoidance effect). The dashed line represents the uniform probability density value (0.159).

3.5.2.1. Memory

Table 3.5 below shows the mean estimates produced for the memory metrics. The model accurately reproduced the pattern of effects observed in Chapter 2, showing greater accessibility in the clustered compared to the non-clustered condition but greater precision in the non-clustered compared to the clustered condition. The estimates derived are much larger than found in our experiments, with most measures being outside of the 95% CIs derived.

Similarly, the size of the differences in conditions is much larger than found in the behavioural experiments.

Table 3.5.

Descriptives for memory variables from Model Family 2, including identification of whether the value was within the 95% CI of the Chapter 2 data and the differences between conditions.

Variable	Condition	*M*	Within 95% CI?	Group Difference
I_p	NC	0.506	N	-0.638
	CL	1.144	N	
I_k	NC	1.495	N	0.275
	CL	1.220	Y	
PKD: Memory	NC	0.045	N	NA

Note. I_p = Accessibility, I_k = Precision. PKD = Kernel density at the cluster centre. NC = Non-Clustered, CL = Clustered. Y = Yes, N = No. Group difference was calculated by subtracting the CL mean from the NC mean. NA = Not applicable.

3.5.2.2. Generalisation

The mean estimates derived for generalisation trials are shown in Table 3.6 below. For all estimates, the mean values were outside the 95% CI derived for Chapter 2 (see Table 3.1 above). Both models generally produced the same behaviour, though the encoding-based model produced lower overall estimates compared to the retrieval-based model; this is also illustrated in Figure 3.3 above.

Table 3.6.

Descriptives for generalisation variables from Model Family 2, including identification of whether the value was within the 95% CI of the Chapter 2 data

Model	Condition	*M*	Within 95% CI?
Encoding	NC	0.085	N
	CL	0.382	N
Retrieval	NC	0.028	N
	CL	0.473	N

Note. NC = Non-Clustered, CL = Clustered. Y = Yes, N = No.

3.5.3. Discussion

The second family of models aimed at examining whether modulating retrieval probability produced an avoidance effect during generalisation. Specifically, making items in the non-clustered condition have a higher probability of retrieval the further from the cluster they appear. Unlike the first family of models, the current set of models produced an avoidance effect during generalisation. However, an avoidance was also present for non-clustered memory trials.

Along with producing an avoidance behaviour, this second family of models reproduced the other pattern of results observed in Chapter 2. Specifically, greater accessibility in the clustered compared to the non-clustered condition but reduced precision in the clustered condition compared to the non-clustered. As in Model Family 1, this was a function of the independent p and k inputs for the clustered and non-clustered conditions. Therefore, it was not a product of the underlying mechanisms of the model itself. Nevertheless, there was clear evidence of generalisation occurring for the clustered condition. Unlike the accessibility and precision differences, the generalisation behaviour observed was a product of the model. Notably, however, these effects were larger than in Chapter 2 (this is discussed in more detail in the General Discussion, below). Therefore, Model Family 2 recreated the general pattern of results observed in our behavioural data and predicted that the avoidance effect should be present during memory and generalisation.

3.6. Memory and the Avoidance Effect

Model Family 2 predicted that the avoidance behaviour presented in generalisation trials was driven by an avoidance in the non-clustered old (memory) trials. To test this prediction, an analysis of the memory data from Chapter 2 and an independent secondary dataset (Berens et al., 2020) was conducted to identify whether there was evidence of avoidance in non-clustered old items. The analyses were undertaken to falsify the previous model, which relied on memory showing evidence of avoidance for it to be present during generalisation. Therefore, if there is no evidence of avoidance in memory, this model can be falsified. However, if there is evidence of avoidance in memory, the model correctly predicted the need for non-clustered old items to show an avoidance for novel items to show this behaviour too.

3.6.1. Exploratory Analysis: Chapter 2

To conduct this analysis, a GLME was computed. The GLME was fit similarly to those described in Chapter 2, with a gamma distribution to model data spread, a log-link function and estimated using the maximum likelihood fitting method within the MATLAB Statistics and Machine Learning Toolbox. The dependent variable was the probability density of non-clustered items at the centre of the experimental imposed von Mises. These were compared to a circular uniform density value ($2\pi^{-1}$). An assessment of whether the effect in memory and generalisation differed in magnitude was also undertaken. Therefore, both memory and generalisation non-clustered trials were included in the analysis. The model included three fixed effects: Trial Type (0 = Memory, 1 = Generalisation), Delay (0 = No Delay, 1 = Delay) and Setting (0 = In-Lab, 1 = Online), along with their interactions. For the model, random intercepts for subject were included along with random slopes for the effect of Trial Type. Both Cohen's *d* and Bayes Factors were computed as described in Chapter 2, only examining the fixed effects. Mean estimates were derived from the model. Though generalisation data was included, this was only used to assess whether the effect in generalisation was larger than memory. All other effects only considered the memory data. Appendix A shows the contrast matrices used to examine the effects of interest.

One consideration to be made is how the 95% confidence intervals discussed in Table 3.1 (above) suggested non-clustered memory items could be larger than the uniform density value (0.159), meaning there was a lack of evidence for avoidance. However, the estimates reported in Table 3.1 were derived based on the Student's *t*-distribution. Though these estimates are generally acceptable to use, we do need to be careful before making formal conclusions about whether an effect is present or not. Specifically, the data collected in Chapter 2 are not normally distributed but instead adhere to a gamma distribution (i.e., cannot go below a value of 0). Consequently, it may not be appropriate to conclude that there was no evidence of avoidance in memory without analysing the data appropriately. Instead, applying a GLME that models the data spread on a gamma distribution will provide a more accurate estimate of whether avoidance is present in memory.

Figure 3.4 shows the mean density and distribution of locations selected by participants for memory trials. It was found that the probability density for non-clustered old words was lower than predicted by a uniform density, $t(514) = 3.01$, $p = .003$, $d = 0.07$, $BF_{01} = 0.43$, suggesting a similar avoidance behaviour for non-clustered old words as was present in non-clustered novel words. Of note, the Bayes Factor remains inconclusive despite the significant effect. No effect of delay ($F(1,514) = 0.01$, $p = .941$, $d < 0.01$, $BF_{01} = 13.71$) or setting ($F(1,514) = 0.29$, $p = .592$, $d = 0.03$, $BF_{01} = 12.23$) were observed when examining only memory data. However, there was an interaction between the two, $F(1,514) = 3.99$, $p = .046$, $d = 0.21$, $BF_{01} = 1.27$. Notably, the Bayes Factor for the interaction effect was inconclusive. Exploration of the post-hoc effects found no significant effects even before correction ($p \geq .078$, $d \leq 0.16$, $BF_{01} \geq 2.25$). Finally, an examination of differences in the non-clustered densities of memory and generalisation items occurred. For this comparison, memory ($M = 0.151$, $SE = 0.02$)

showed significantly less avoidance than generalisation (*M* = 0.136, *SE* = 0.03), $F(1,514) = 28.62$, $p < .001$, $d = 0.23$, $BF_{01} = 1.41 \times 10^{-5}$.

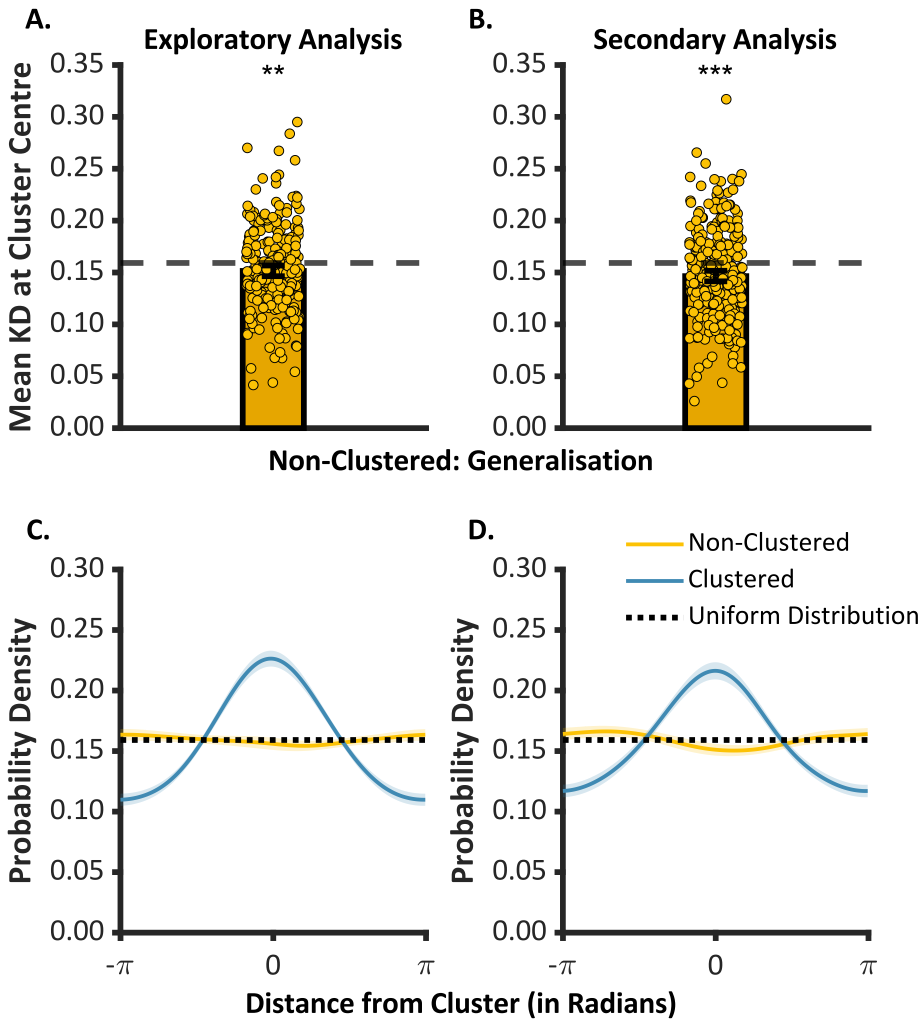

Figure 3.4. **Locations selected for memory trials**. *A-B*: Mean Probability Density for Memory trials for (A) Chapter 2 data, or (B) the secondary analysis. Individual data points represent participant scores. The dashed line represents the uniform probability value (0.159). *C-D*: Spatial distribution of locations selected for old words centred to the experimentally imposed von Mises distribution for (C) Chapter 2 data, and (D) secondary data. Data are collapsed across all delays. Error bars represent 95% confidence intervals around the mean for all plots. ** = $p < .01$, *** = $p < .001$.

The above suggests the presence of a similar avoidance effect in memory, as was shown in generalisation. Specifically, participants show a reduced probability density at the centre of the cluster for non-clustered items. However, the effect in memory is significantly smaller than what was observed in generalisation. Nevertheless, these results suggest a bias in memory and supports the prediction of Model Family 2 (above), which indicated that a global non-clustered avoidance behaviour should be present. Still, caution should be taken before firm conclusions are drawn from the present analysis as it was exploratory and the Bayes Factor remained anecdotal. To overcome this, analysis on an independent secondary dataset occurred; this was pre-registered before any analysis took place. This secondary dataset used a similar methodology but did not include generalisation trials. The avoidance effect could be viewed as a robust phenomenon if the same avoidance behaviour was observed in the independent dataset.

3.6.2. Confirmatory Analysis

The data collected for the Berens et al. (2020) registered report was explored. Briefly, in Berens et al. (2020), they used the precision paradigm to explore how forgetting differentially impacted accessibility (i.e., proportion remembered) and precision (i.e., degree of error from target location to response location). It was found that the presence of a schema (or pattern) benefited levels of accessibility positively but to the detriment of precision. As this experiment was conducted using a similar paradigm, an exploration of whether there was evidence of an avoidance effect was undertaken; this would act as a confirmatory analysis of our previous exploratory analysis. Consequently, if evidence of an avoidance effect was found, firmer conclusions can be drawn related to this effect. The pre-registration for this secondary analysis can be found here: https://osf.io/pwy5t/.

3.6.2.1. Hypotheses

It was predicted that participants' non-clustered responses would show evidence of avoidance, which would not be affected by delay. Specifically, participants would show a reduced probability density at the centre of the experimentally imposed cluster for non-clustered items. This was to be evidenced by a significant reduction from the uniform probability density value.

3.6.2.2. Exclusion Criteria

The only data that were explicitly excluded were study-test delays greater than 24 hours. Beyond that, the same exclusion criteria were applied to the data as was reported in the previous experimental Chapter (see Chapter 2: Exclusion Criteria). Briefly, individual trial exclusions were applied when: (1) study trials were repeated five or more times, and (2) no responses were given on a test trial. Datasets were excluded when: (1) the number of test trials timed out more than 20 times for clustered or non-clustered items, and (2) the data did not converge using mixture modelling.

3.6.2.3. Participants

The dataset contained 431 participants (63% female), with a mean age of 27.23 years (*SD* = 5.07). There were 401 usable datasets (64% female) following data cleaning with a mean age of 27.25 years (*SD* = 5.06). Thirty participants were excluded because: 17 had incomplete datasets at Test due to inattention, eight did not respond to 20 or more trials, and five datasets did not converge using the mixture model. Notably, these 401 participants included all delay conditions (0 hours up to 96 hours). Once both 48-hour and 96-hour delays were removed, the final sample consisted of 294 participants (62% female, M_{Age} = 27.12 years, SD_{Age} = 5.05), with generally equal numbers across delays (0hrs = 57, 3hrs = 63, 6hrs = 60, 12hrs = 58, 24hrs = 56).

3.6.2.4. Analyses

The GLME for these analyses was fit using the same parameters as described for the exploratory analysis above. However, the model only included a main effect of Delay; this was dummy coded such that 0hrs delay was the reference variable, and all other delays (3, 6, 12, and 24hrs) explicitly included in the model. The contrast matrices used for these analyses are shown in Appendix A. As there was a directional hypothesis, a one-tailed test was used when exploring the avoidance effect. Further, as has been done previously, Bayes Factors were computed using the same method reported in Chapter 2. However, due to the presence of multiple groups in some instances, an approximation of omega squared (see 2, below) using the *F*-statistic and corresponding degrees of freedom (Albers & Lakens, 2018) was used as a measure of effect size in some instances. When only one numerator degree of freedom existed, the Cohen's *d* was reported and calculated as described in Chapter 2.

$$\omega^2 = \frac{F - 1}{F + \frac{df_{Error} + 1}{df_{Between}}} \tag{2}$$

The distribution of locations selected by participants collapsed across delay is shown in Figure 3.4 (above). As can be seen, there appears to be some evidence of avoidance behaviour present in the locations selected by participants. First, differences in the probability density estimates resulting from the effect of delay were examined. No significant main effect of delay was found, $F(4,289) = 0.15$, $p = .963$, $\omega^2 = -.01$, $BF_{01} = 364.51$. Consequently, the probability densities do not appear to change as a function of time; this is similar to the exploratory analysis of the Chapter 2 data. Next, evidence of avoidance was explored. It was found that, irrespective of delay, participants showed significant avoidance of placing locations at the centre of the cluster for non-clustered items, $t(289) = 4.71$, $p < .001$, $d = 0.13$, $BF_{01} = 5.03 \times 10^{-4}$. Again, this replicates the previous exploratory analysis and shows that the avoidance effect is present in memory. However, the evidence in favour of this conclusion was more substantial than found in the exploratory analysis.

3.6.3. Discussion

These analyses aimed to determine whether the second set of models reported in this chapter (i.e., Model Family 2) had correctly predicted the presence of an avoidance behaviour in non-clustered memory (old) trials. In line with this prediction, the exploratory and confirmatory analyses found evidence of avoidance during memory. This finding supports the predictions of the second set of models, showing that for an avoidance effect to be present during generalisation, there needs to be avoidance in memory. Along with this finding, no evidence of a change over time was identified, suggesting the avoidance effect remained even after a 24-hr period; this is identical to generalisation trials. Additionally, the avoidance observed during memory was significantly less than observed during generalisation.

In Model 2, retrieval-based generalisation was found to produce greater avoidance behaviour compared to memory avoidance. In support of this prediction, the present analysis found that the avoidance behaviour in memory was smaller than found in generalisation. One reason for this may relate to items with higher retrieval probabilities (i.e., those on the opposite side of the circle) being more easily accessible when attempting to generalise. Whilst old items can rely on either remembering the old item or generalising, novel items can only rely on generalisation processes. As a result, given that items further from the cluster may be more accessible, the avoidance effect may be exacerbated during generalisation.

In contrast, the encoding-based model found the avoidance behaviour was smaller during the non-clustered novel trials than old (memory) trials. One reason for this may be that the encoding-based model used all remembered items, irrespective of their retrieval probability, when generating parameter estimates for the von Mises distribution. The estimated parameters were then used to generate a series of “random” locations (adhering to the parameters specified). As such, using all remembered responses as opposed to a subset produced lower levels of avoidance in the non-clustered condition for this model type.

In conclusion, Model Family 2 provided an accurate account for the behaviour observed in Chapter 2 and generated a novel prediction related to non-clustered memory trials. Specifically, it was predicted that memory trials would show a similar avoidance behaviour to generalisation. In line with this prediction, both the exploratory and confirmatory analyses found that non-clustered memory items showed a reduced density in the number of locations selected by participants at the cluster centre. These results support the second family of models by demonstrating that avoidance was present in both memory and generalisation.

3.7. Model 3: Interference

Many studies have used the precision paradigm as a means of investigating schema use in memory (Antony et al., 2021; Berens et al., 2020; Harlow & Donaldson, 2013; Richter et al., 2019) and generalisation (Tompary et al., 2020). However, in many of these studies (e.g., Antony et al., 2021; Berens et al., 2020; Tompary et al., 2020), the presence of a “schema” appears to decline as memory itself declines. These findings raise an important consideration for the conclusions made: Is what is being investigated a function of schema development or some other phenomenon? Schemas are believed to be long-lasting and should not decay as memories for individual events do (Ghosh & Gilboa, 2014; van Kesteren et al., 2012). Therefore, findings of decay over time may suggest an alternative mechanism is involved. One mechanism proposed here is the influence of interference.

The present model aimed to assess whether an interference mechanism could explain the behaviours observed in Chapter 2. Whereas the second family of models modulated retrieval probability based on the target location (something the present model will also implement), it did not propose a specific mechanism that led to this effect (i.e., it resulted from a retrieval probability function). In contrast, the present model proposes a mechanism for explaining these changes in retrieval probability (i.e., proximity-based interference). Developing this model may shed light on the mechanisms driving the avoidance effect, thus extending the second family of models.

For the present model, no p or k parameters were explicitly included, no differentiation between clustered and non-clustered items was made, and both memory and generalisation trials were treated similarly. Further, interference occurred at both encoding (proximity-based interference) and retrieval (semantic-based interference). This hybrid-based approach was achieved by affecting retrieval probability at encoding based on the proximity of previously presented items. Then, a sample of items would be selected at retrieval based on semantic proximity to the target word. The retrieval probability and semantic distance were used as weights to affect the mean location selected for a given trial.

3.7.1. Method

Semantic distance was incorporated into the model using the 240 words from Chapter 2. As a reminder, there were two-word groups: human-made (e.g., calculator, hammer) and natural (e.g., apple, desert). The semantic distance between words was estimated via Euclidian distance and derived from the pre-trained word2vec model (Mikolov et al., 2013). Words were selected to optimise semantic similarity within a category and distance between categories. In other words, the distance between words of the same category (e.g., human-made) was small, whilst the distance between categories, in this case, the distance between human-made and natural, was large. For the model, these distances were rescaled to between 0 and 1, with larger values indicating greater similarity between words. The rescaling was done so both retrieval probability and semantic distance were on the same scale (i.e., bounded between 0 and 1). In this way, they could be used as weights during location selection (see Retrieval, below).

3.7.1.1. Experimental Parameters

Similar to the previous models, the present model was provided with target locations for items. However, the input also provided information about the order of presentation and which word was presented on each trial. Knowledge of the order in which encoding took place and which item was presented were essential for introducing interference at both encoding and retrieval, respectively. The order of presentation and stimuli used were randomised for each "participant", similar to the experiments described in Chapter 2.

3.7.1.1.1. Encoding

During encoding, the model modulated the probability that an item would be later retrieved based on its proximity to previously encoded items (irrespective of the clustering condition). Each "participant" had a maximum retrieval probability; this was a random value between 0 and 1. The maximum retrieval probability initially varied across participants. However, varying the initial value had little to no impact on the final weighting (i.e., combining proximity- and semantic-based interference) as the relative difference between weightings remained the same regardless of initial probability. This is discussed in more detail in the Retrieval section, below. Nevertheless, during encoding, this retrieval probability value was affected when an item was proximal to other previously encoded items. The present model's proximity value was set to 90° (i.e., 45° on either side of the newly encoded item); this was selected as the degree of proximity did not drastically affect model behaviour. An example of the proximity function is shown in Figure 3.5 below. If the new item was proximal to any previously encoded items, the initial probability of remembering was divided by the number of proximal items (regardless of condition). If no items or only one item were proximal, the probability value would be unchanged. Once all items had been encoded, the model then moved to the retrieval phase.

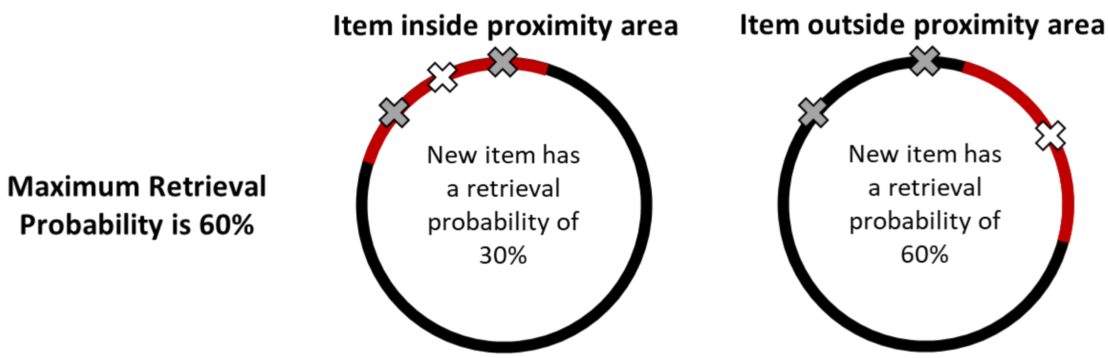

Figure 3.5. **Illustration of the proximity function**. A maximum retrieval probability value was set for each participant; this value determined the probability that an item would be retrieved from memory. During encoding, if an item were located within 90° of any previously encoded items (grey crosses), this new item (white cross) would have a reduced probability of retrieval. The red region represents the "proximity" function illustrating the region where an item would need to appear to affect retrieval probability. This new probability value was determined by dividing the maximum retrieval probability by the number of proximal items. For example, if the initial proximity value was 60% and the new item was at a location with two other items, its retrieval probability would be 30%. However, if the new item was not proximal to any previously encoded item, the probability of retrieval would be unchanged.

3.7.1.1.2. Retrieval

During retrieval, locations were selected for both memory and generalisation trials. Here semantic-based interference was introduced; this occurred as the target item would also activate memory for locations of close semantic neighbours (i.e., there was a spreading of activation). On each trial, the model selected semantically similar items to the present trial; these items would have previously been associated with a location (i.e., presented during encoding). For memory trials, the selected items included the target itself and two other semantically related items. For example, if the target item were "calculator", locations for "smartphone" and "robot" would be retrieved alongside "calculator", given they were close semantically to the target item. For generalisation, three semantically related items were retrieved. For instance, had the word been "desert", items such as "sand", "snow", and "cactus" were retrieved as there would be no location associated with the word "desert". Apart from this selection of words, the memory and generalisation trials were identical.

The retrieval of these other locations acted as a form of semantic-based interference as they influenced the location judgement made by the model. The mean location was computed from the retrieved locations but weighted by retrieval probability (proximity-based interference) and semantic distance (semantic-based interference). Specifically, the retrieval probability value for each item and their corresponding semantic distances were multiplied to form a weighting for each item when computing the mean location. In the case of memory items, the semantic distance weighting would favour the original word such that "calculator" was scored 1, whilst all other items varied from 0 to 1 based on semantic similarity. Using this weighting allowed the model to account for how biases in reporting may arise due to interference on the location and semantic level. Note, this model does not distinguish between "remembered" and "non-remembered", and no additional precision "noise" is added to the target locations (c.f. Model Family 1 and 2).

Returning to the initial retrieval probability value, it is clear that the relative differences would remain the same regardless of the initial value. For instance, if the retrieved angles were 150°, 180° and 210°, with retrieval probabilities of 0.7, 0.7 and 0.35, and semantic distances of 0.5, 1.0 and 0.2, the weighted mean angle would be 172.50°. However, had the initial probability value been 0.5 rather than 0.7, the weighting function would still produce the same weight mean angle of 172.50°. Therefore, the initial value itself had little influence on the model, but (as discussed below) the retrieval probability was necessary.

3.7.2. Results

Figure 3.6, below, shows the average locations selected by the model for both memory and generalisation trials. As shown, the model appeared to have generated an avoidance behaviour in both memory and generalisation, along with showing a clustering behaviour for the clustered condition across memory and generalisation.

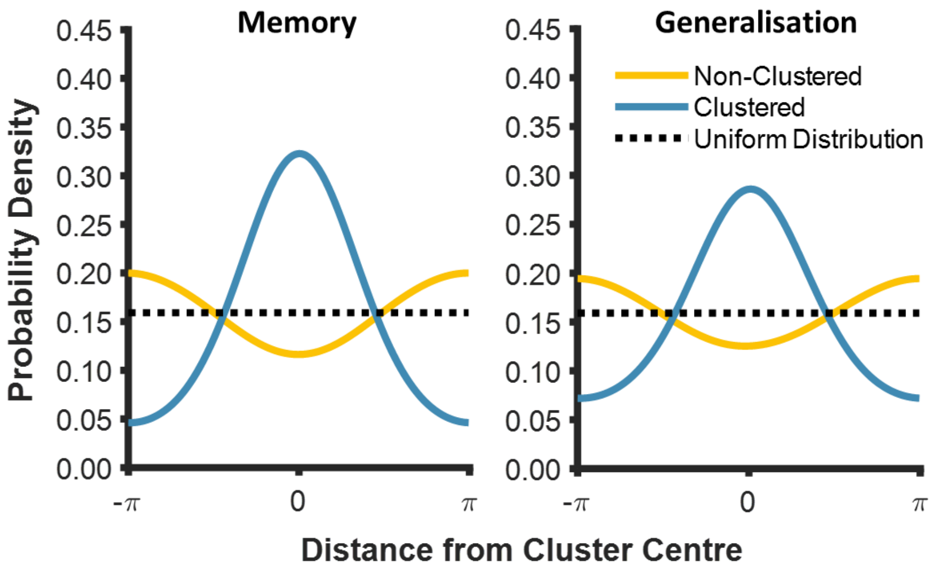

Figure 3.6. **Model 3: Density of Locations**. Plots show the distribution of locations selected by the model for both memory and generalisation items, centred to the experimentally imposed von Mises distribution. The dashed line represents the uniform probability density value (0.159).

Examination of the mean estimates (see Table 3.7, below) shows that the model correctly mapped the broad pattern of results present in Chapter 2. First, the model was able to show an accessibility benefit in the clustered condition compared to the non-clustered. Additionally, the non-clustered benefit in precision was also found. Unlike the previous models where p and k parameters were used to control behaviour, these differences were a consequence of the workings of the model. Along with this, the model produced a generalisation behaviour for clustered items showing greater kernel density for clustered items compared to the uniform value. Finally, the model correctly generated an avoidance behaviour for both memory and generalisation.

Table 3.7.

Model 3 descriptives for all variables, including whether the value was within the 95% CI of the Chapter 2 data and the differences between conditions.

Variable	Condition	*M*	Within 95% CI?	Group Difference
I_p	NC	1.091	N	-0.115
	CL	1.206	N	
I_k	NC	1.641	N	0.037

	CL	1.604	N	
PKD: Memory	NC	0.116	N	NA
PKD: Generalisation	NC	0.125	N	NA
	CL	0.286	N	NA

Note. I_p = Accessibility, I_k = Precision. PKD: Memory = Kernel density at cluster centre for memory trials. NC = Non-Clustered, CL = Clustered. Y = Yes, N = No. Group difference was calculated by subtracting the CL mean from the NC mean. NA = Not applicable.

The size of the effects reported differ from those in Chapter 2, with all values being outside the 95% confidence intervals described (see Table 3.1 above). In this instance, the accessibility effect was larger, and the precision effect was smaller than in Chapter 2. Further, the density of locations for clustered generalised items and non-clustered memory and generalisation items was larger than in Chapter 2. For the avoidance behaviour, memory showed greater avoidance compared to generalisation, contrary to the earlier exploratory analyses. However, the difference between memory and generalisation from the model output was small (0.009). Therefore, the models' slight inconsistency may not be too problematic. However, further analysis may be required. Despite these apparent differences, the model correctly maps many of the patterns observed in Chapter 2 and the analyses reported in this Chapter.

3.7.2.1. Importance of Proximity and Semantic Distance

An open question was whether both proximity-based interference and semantic-based interference were required to produce the effects reported above. To explore this question, two analyses were conducted. The first analysis removed proximity-based interference at encoding, with retrieval probabilities not influenced by the proximity of items. Instead, retrieval probabilities remained the same for all items. However, semantic-based interference remained, with semantic neighbours to the target still being retrieved and used when making a location judgement. The second analysis removed semantic-based interference. Specifically, three randomly selected items would be retrieved rather than close semantic neighbours, meaning semantic proximity did not affect location judgements. However, retrieval probabilities were still manipulated at encoding via proximity-based interference. These retrieval probabilities were calculated based on proximity-based interference and used as weights when sampling from memory, such that items with higher retrieval probabilities were more likely to be used when deciding the location of an item (similar to Model Family 2). Figure 3.7, below, shows the effects of removing proximity-based and semantic-based interference from the models.

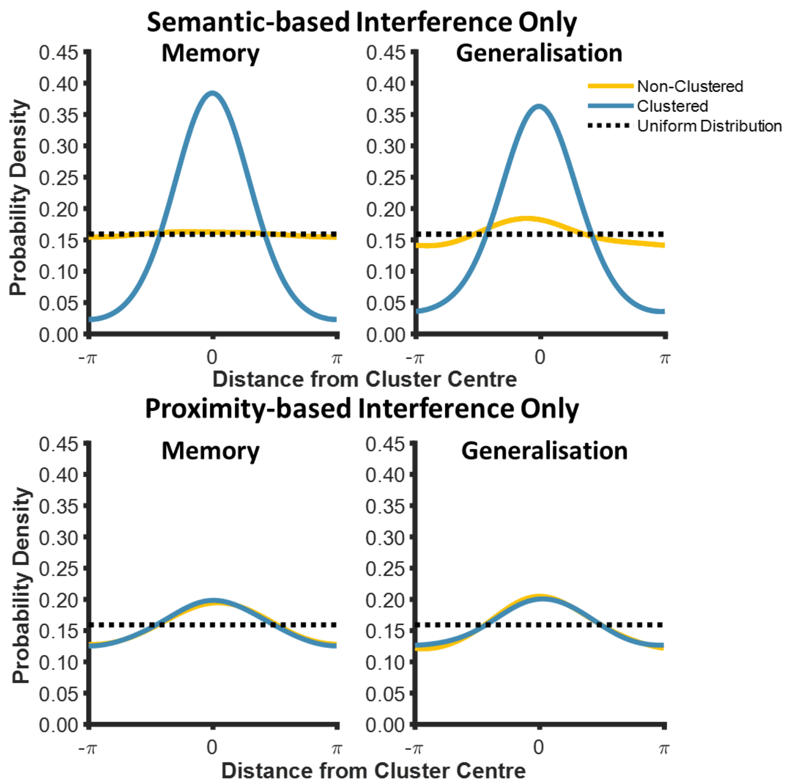

Figure 3.7. **Removal of different forms of interference from the model**. The plots show the distribution of locations selected for memory and generalisation items, centred to the experimentally imposed von Mises distribution. The top row shows how the presence of only semantic-based interference affects the model output. The bottom row shows how the presence of only proximity-based interference affects the model output. The dashed line represents the uniform probability density value (0.159).

3.7.2.1.1. Removal of Proximity-based Interference (Encoding)

As shown in Figure 3.7, when locations are derived based only on semantic-based interference, the clustered condition showed increased density towards the cluster centre, with some evidence of non-clustered items being treated more uniformly. The results suggest that removing proximity-based interference still produced the pattern of effects in accessibility, with greater accessibility in the clustered (*M* = 1.693) compared to the non-clustered (*M* = 1.578) condition. However, contrary to the findings of Chapter 2, precision was greater in the clustered (*M* = 1.524) compared to the non-clustered (*M* = 1.409) condition. All values were larger than reported in Chapter 2 and outside of the 95% confidence intervals. Further, the lack of proximity-based interference resulted in both non-clustered conditions showing evidence of peaks within the cluster centre, with non-clustered memory (*M* = 0.163), generalisation (*M* = 0.182) and clustered generalisation (*M* = 0.363), all possessing values larger than a uniform density (0.159). However, non-clustered memory is within proximity to the uniform value. Therefore, semantic-based interference alone is not sufficient to produce the pattern of behaviour from Chapter 2.

3.7.2.1.2. Removal of Semantic-based Interference (Retrieval)

When semantic-based interference was removed from the model (i.e., random words were selected), greater accessibility in the clustered (*M* = 0.454) than in the non-clustered (*M* = 0.428) condition remained. Similarly, precision was greater in the non-clustered (*M* = 1.292) compared to the clustered (*M* = 0.999) condition. These two results are in line with the pattern of behaviour observed in Chapter 2. However, the values are outside of the 95% confidence interval range, being lower than expected. Interestingly, as shown in Figure 3.7, the model treated both the clustered and non-clustered conditions similarly. Specifically, the two conditions showed similar density at the cluster centre for memory and generalisation.

Examination of the mean estimates found non-clustered memory (*M* = 0.194), non-clustered generalisation (*M* = 0.205) and clustered generalisation (*M* = 0.201) were all greater than the uniform density value (0.159) and were similar in magnitude. Again, this supports the notion that both proximity- and semantic-based interference need to work together to produce the pattern of effects observed in Chapter 2.

3.7.3. Discussion

The present model aimed to assess whether an interference mechanism could predict the pattern of results observed in Chapter 2. As demonstrated, the present model could remap the behaviour patterns for accessibility, precision, generalisation, and avoidance. However, the effect of avoidance was larger in memory than generalisation, contrary to the exploratory analysis reported earlier. Nevertheless, it appears that interference at both encoding (i.e., proximity-based interference) and retrieval (i.e., semantic-based interference) was sufficient to produce the effects of interest without explicit separation of clustered and non-clustered conditions or inclusion of p and k parameters.

In two exploratory analyses, it was shown that both forms of interference were necessary to produce the effects of interest. Specifically, removal of proximity-based interference resulted in precision showing the opposite pattern of effects along with a lack of evidence for avoidance. In contrast, removing semantic-based interference resulted in a lack of differentiation between clustered and non-clustered items.

One criticism that could be raised with the interference model is that it only considers proactive interference without considering the effects of retroactive interference (e.g., Kliegl et al., 2015). Though true, it is unlikely that implementing a retroactive interference mechanism would necessarily change the model output. If interference during encoding were to favour items towards the end of the encoding session, it would still lead to reduced retrieval probability for items closer together. Consequently, interference would continue as more items appear within the clustered area. Thus, a similar set of behaviour may be observed given that there would still be a large proportion of clustered, compared to non-clustered trials appearing within one area of the circle. Consequently, recall of these items would be suppressed by the abundance of clustered items, thus resulting in an avoidance behaviour. However, this is speculative, and it would be worth examining whether these two forms of interference have differential influences on behaviour. This could be explored using the model itself or examination of the behavioural data obtained in Chapter 2. An analysis controlling for encoding order could potentially suggest the presence of an interference-effect. If proactive interference occurred, items first encoded should be more accurately reported (i.e., show less angular error) than those encoded later. In contrast, for retroactive interference, items encountered at the end of learning should be more accurately recalled than those that appeared earlier.

In summary, the interference model appears to provide a more parsimonious explanation for the behaviour observed in Chapter 2. Unlike the previous models, which required an explicit dichotomy between clustered conditions and the use of p and k parameters to model differences in accessibility and precision, the interference model produced these effects based on proximity- and semantic-based interference alone. Further, it was able to show the ability to generalise to novel instances and produce an avoidance effect for non-clustered items.

3.8. General Discussion

This Chapter aimed to explore mechanisms for the avoidance behaviour observed in Chapter 2. The first family of models assessed whether basic forms of encoding- and retrieval-based models would predict the presence of the avoidance behaviour. Neither model produced the avoidance effect. In Model Family 2, where memory was modulated based on the proximity of items to the cluster, an avoidance effect was predicted in generalisation. However, the model also made a novel prediction – for avoidance to be present during generalisation, it should also be present during memory. Examination of this hypothesis through exploratory (using Chapter 2 data) and confirmatory (using Berens et al., 2020 data) analyses supported the model's prediction. A final model was then developed to provide a more mechanistic view on why the avoidance effect may occur. Specifically, whether an interference mechanism alone could predict the avoidance effect. This final model predicted the avoidance behaviour, matched the broad behavioural patterns reported in Chapter 2 and provided a more parsimonious solution than the previous two models as to why the avoidance behaviour was present.

Finding that neither an encoding- nor retrieval-based model could predict avoidance was not surprising. As discussed in the Introduction, both model types would reproduce the pattern presented during encoding at retrieval. For clustered items, participants should be more likely to place locations within the clustered area. In contrast, for non-clustered items, locations should be placed without bias across the entirety of the circle. This is what was shown by Model Family 1, where there was evidence of a peak in probability density for clustered items towards the centre of the cluster, with non-clustered items showing little to no difference from the uniform density.

Model Family 2 examined how modulating retrieval probability as a function of proximity to the cluster affected generalisation. This modulation process was similar to how schema may modulate memory. Specifically, a schema may increase the probability of remembering an item when it is congruent with expectations but decrease probability when items are not congruent with those expectations (Brewer & Treyens, 1981; though see Greve et al., 2019 for evidence of incongruency leading to similar levels of memory as schema-congruent). Non-clustered items appearing within the clustered area of the circle may act as incongruent items, given that they may violate expectations. This model demonstrated an avoidance behaviour during generalisation but required the presence of the same behaviour during memory. This novel prediction was supported by both an exploratory and confirmatory analysis. Discussion of what this may mean for the broader literature are covered in Chapter 5.

The final model expanded on the second family of models by exploring possible mechanisms (i.e., interference) that could explain the presence of the avoidance effect. During encoding, items were given lower retrieval probabilities when they were close to previously encoded items. At retrieval, target items were retrieved, along with semantic neighbours, and their average was taken to generate a location estimate on all trials. These mechanisms are similar to other interference-based proposals (e.g., spreading activation; Anderson & Pirolli, 1984, or cue overload; Watkins & Watkins, 1975). The model produced greater accessibility in the clustered condition compared to the non-clustered, but reduced precision in the former compared to the latter condition. There was evidence of generalisation for clustered items and avoidance behaviour for non-clustered items during memory and generalisation. Unlike the previous models, this final model only required input related to the target locations, without the need for explicit coding of conditions (i.e., clustered vs non-clustered) or the use of parameters (i.e., p and k) to modulate behaviour in memory. As a result, this model provided a parsimonious explanation for why avoidance was present within the behavioural data. How this model may also apply to other literature (e.g., Tompary et al., 2020) is discussed in Chapter 5. Nonetheless, given the lack of statistical comparisons conducted between the model and real-world data, some caution should be taken with interpreting the results reported.

The inclusion of this type of spreading activation (activation of competing memories) is comparable to the Target Confusability Competition (TCC) model developed by Schurgin et al. (2020). Their model uses psychophysical scaling of similarity and signal detection to derive behavioural responses for working and long-term memory precision tasks. If you consider the example of learning about colours in a circular environment, when asked to report the colour you were presented with, this may result in the activation of other related colours. This spreading of activation relates to the model's psychophysical proximity parameter, which identifies closely related colours and activates those representations during recall. For instance, if you were presented with green, activation of near neighbours such as blue and yellow may occur, which introduces noise to the system. It is then up to the participant to determine the correct signal. This signal strength is based on *d'* (i.e., how strong is the familiarity-based signal). On average, the colour observed will be the colour chosen, but the noise introduced via familiar psychophysical signals may mean an unseen colour (e.g., yellow) is selected in some instances. This model has parallels to the interference mechanism proposed here. Specifically, using semantic-based interference to modulate location judgements is similar to their psychophysical scaling mechanism. Future work should explore which model better fits the plethora of findings across working and long-term memory paradigms or whether the inclusion of location- and semantic-based interference in the signal detection framework of the TCC model can explain the present results.

Though three models were produced for the present chapter, it is worth noting that these are not the only possible explanations for the avoidance effect. As discussed in Chapter 2, both a mutual exclusivity bias (Clark, 1988) or base rate neglect (Hawkins et al., 2015) may explain the avoidance effect. For instance, base rate neglect proposes that participants use relative rather than absolute probabilities when making location judgements. If location judgements were based on absolute probabilities, participants would remap the distributions accurately (i.e., select locations uniformly for non-clustered items but follow a von Mises for clustered items). However, as non-clustered items differed in density compared to clustered items in one area of the circle, the relative probability of non-clustered items was lower in that region compared to the other side of the circle. Therefore, location responses would show the avoidance behaviour presented in Chapter 2. In future, it may be useful to discern which of the proposed mechanisms (interference, mutual exclusivity, or base rate neglect) provide the best fit for the behaviour observed in Chapter 2. This is discussed further in the final chapter of the thesis.

Another point of consideration is how the models often overestimated the size of the effects; this was particularly evident for the second and third models. Speculatively, the reason for this overestimation may relate to underlying assumptions made within the models. At present, the models produce behaviour based on the assumption that participants extracted the pattern and used this at retrieval without any random guessing. The debrief data from Chapter 2 suggests that around 42% (109/261) of participants did not perceive a pattern. The lack of knowledge about the pattern may have resulted in worse performance as participants turned to “random guessing” for forgotten or novel trials (though this is speculative). Ignoring this possibility in the models may have resulted in an overestimation of behavioural performance. Therefore, the addition of random guessing on some trials may be required to produce more representative behaviour. Future work may wish to consider analysing the effects of interest, isolated to whether the participant reported extracting the underlying pattern.

Overall, the present Chapter aimed to assess possible explanations for the avoidance effect in generalisation. All three models were able to remap a similar pattern of behaviour observed for measures of accessibility and precision. For the first two models, this was produced by directly controlling these metrics within the model. For the interference model, these differences in accessibility and precision occurred without explicit coding. Though the first set of models did not predict an avoidance effect, the final two models did. For the second model, a novel prediction about the avoidance behaviour was made. Specifically, the presence of avoidance behaviour in memory was proposed; this was subsequently found via exploratory and confirmatory analyses. This behaviour was also observed in the final (interference) model. The interference model produced the behaviours of interest with fewer free parameters and provided a parsimonious explanation for why an avoidance behaviour may have been present in the data.

Chapter 4:
Preliminary Results: fMRI-based investigation of Memory and Generalisation

4.1. Abstract

The present work aimed to examine the neural correlates of memory-based generalisation using the precision task developed in Chapter 2. Due to the COVID-19 pandemic, a preliminary analysis of an incomplete dataset and experimental plan are presented. First, a behavioural pilot was conducted to investigate whether completing a semantic categorisation task (SCT) before the precision paradigm impacted the behaviour observed by comparing the results to those found in Chapter 2. The SCT was implemented to minimise item novelty effects present during generalisation trials. During the SCT, participants were presented with all 240 words and categorised them as either human-made or natural. Following this, they completed the study and test phase of the precision paradigm. No evidence of significant changes resulting from the SCT's inclusion were found, with the same pattern of effects present as described in Chapter 2. Therefore, the SCT was implemented into the fMRI investigation. In the fMRI pilot, participants completed the SCT and study phase of the precision paradigm outside the scanner. In the scanner, participants completed the test phase. The results showed significantly greater BOLD activity within the vmPFC during memory compared to generalisation trials. Further, for the same contrast, there was marginally significant activation within the dorsal striatum. Speculations as to why the vmPFC and dorsal striatum were active are discussed. The chapter ends considering possible design (e.g., including subjective judgements) and analysis (e.g., functional connectivity analysis) changes along with open questions that may be useful for an independent investigation.

Keywords: *hippocampus, ventromedial prefrontal cortex, generalisation*

4.2. Introduction

The medial temporal lobe (MTL) and prefrontal cortex (PFC) are said to have critical roles in memory-based generalisation, particularly in how they interact (Kumaran et al., 2009; Liu et al., 2016; Preston et al., 2004; Raykov et al., 2020; Shohamy & Wagner, 2008; van Kesteren et al., 2012; Zeithamova et al., 2008). Whilst the MTL allows for rapid binding of memory representations, the PFC plays a role in extracting regularities across experiences to form schema. These schema represent the central tendencies across events (e.g., the average) lacking any unique episodic details (Kroes & Fernández, 2012; Preston & Eichenbaum, 2013).

Studies have shown that participants can generalise their episodic experiences almost immediately post-encoding (Preston et al., 2004; Shohamy & Wagner, 2008; Sweegers & Talamini, 2014; Zeithamova et al., 2012; Zeithamova & Preston, 2010). According to the REMERGE (Kumaran & McClelland, 2012) model, the hippocampus can generalise through retrieval-based means. To achieve this, the hippocampus will act as its own input during learning via big-loop recurrence allowing for any output to become a new input, then reactivating related experiences. Therefore, as an experience is encoded, other related information will become active; this mnemonic information is then sent as an output to neocortical areas and acts as a new input for the hippocampus. Consequently, when generalising, the presence of a partial cue can reactivate all associated items and allow for inferences to be made about a novel, though related, item. Through such a mechanism, the hippocampus can allow for retrieval-based generalisation. Therefore, during the present work, hippocampal activation may be observed during both memory and generalisation.

According to the SLIMMs (van Kesteren et al., 2012) framework, the mPFC and MTL serve different, though complementary, roles during learning. According to this model, activity within the mPFC and MTL is dependent on the congruency of incoming information with pre-existing knowledge (or schema). If incoming information is congruent with a pre-existing schema, this will result in activation of the mPFC. The mPFC will then inhibit activity within the hippocampus as it merges the current experience with pre-existing knowledge. This inhibitory mechanism prevents the MTL from creating an independent experience with no connections to other related events. Instead, the mPFC can integrate this information with existing schematic representations. In contrast, as information becomes increasingly incongruent, the mPFC will have a less inhibitory influence on the MTL and allow the information to be encoded as an individual experience. In this way, the mPFC allows the brain to integrate new information rapidly, in line with more recent work (e.g., Sharon et al., 2011) and theories (e.g., Antony et al., 2017), without the need for MTL encoding processes based on information congruency.

In support of the SLIMMs proposal of congruency influencing vmPFC and hippocampal involvement during learning, van Kesteren et al. (2013) had participants identify whether an object and scene pairing were congruent with expectations (e.g., umbrella and tennis court) by providing subjective ratings of congruency during scanning. These subjective ratings were used to classify object-scene associations as either incongruent, congruent or neither for a given participant. Twenty-four hours post-scanning, participants were presented with both old and new objects and asked to identify if they were old, and if so, what scene it was paired with. It was found that as subjective ratings of congruency increased, so did levels of mPFC activity during encoding, whilst hippocampal activity was greater the more incongruent the stimulus was. As such, the study provides support for the proposals of SLIMMs that hippocampal and vmPFC activation during learning is related to perceived semantic congruency. This may also suggest that differences in activation within the vmPFC and hippocampus may be observed in the present paradigm. For the clustered condition, greater vmPFC activation may be observed due to the potential for a schema to develop. In contrast, the non-clustered condition may show more hippocampal activation as no schema was developed.

The present study investigated the neural correlates of memory-based generalisation using the precision paradigm developed in Chapter 2. In brief, participants learned word-location associations around a circle, with words belonging to two semantic groups. One set of words had a pattern underlying the locations associated with them, whilst the other did not. Using this paradigm, both memory and generalisation can be investigated simultaneously by assessing the patterns extracted by participants moving beyond typical binary outcome measures of correct and incorrect typically used to investigate memory (e.g., Preston et al., 2004) and generalisation (e.g., Bowman & Zeithamova, 2018).

To date, studies that have implemented the precision paradigm within an fMRI setting have focused on memory (Cooper & Ritchey, 2020; Korkki et al., 2021; Richter et al., 2016). For instance, Richter et al. (2016) investigated differences in activation as a function of retrieval success (accessibility), precision and subjective judgements of vividness. Participants encoded and recalled three features (colour, orientation and location) of objects within circular space. All three metrics (accessibility, precision and vividness) could be dissociated when assessing neural activation: accessibility with hippocampal, precision with angular gyrus and vividness with precuneus activation. Therefore, the present work will extend previous research and assess both memory and generalisation under the same conditions. Doing this provides a basis for revealing the neural correlates of memory-based generalisation and may allude to whether encoding or retrieval-based mechanisms are involved in the behaviour observed previously. By examining the preliminary results of this study, any design flaws or open questions for an independent investigation can be identified.

Before presenting the fMRI pilot, a behavioural pilot is presented; this was conducted to assess how the inclusion of a semantic categorisation task (SCT) before learning influenced behaviour in the precision paradigm. The SCT was implemented to address issues related to hippocampal activation associated with item novelty (described in more detail below). As described in Chapter 3, participants rarely acknowledged the superordinate semantic categories (i.e., human-made and natural). Therefore, a pilot experiment was conducted to ensure that explicit awareness of these categories prior to learning did not influence behaviour on the precision task.

4.3. Behavioural Pilot

4.3.1. Introduction

Research has implicated the medial temporal lobe not only in long-term memory but also novelty detection (e.g., Barbeau et al., 2017; Brown & Aggleton, 2001; Grill-Spector et al., 2006; Halgren et al., 1995; Hasselmo & Stern, 2006; Kafkas & Montaldi, 2014; Knight, 1996; Ranganath & Rainer, 2003; Strange et al., 2005). In their study, Strange et al. (2005) had participants categorise items as belonging to one of two groups. This was done through repeated exposure to an object, where participants were given feedback as to whether they classified the object correctly or incorrectly. During test, participants were presented with old and new stimuli and asked to classify them. It was found that novel stimuli (relative to old stimuli) engaged right perirhinal cortex, with anterior hippocampal areas active during learning and decreasing in activity as performance improved due to repeated presentation of a stimulus (i.e., a loss of novelty).

Given that the hippocampus is a region of interest for the fMRI study, any potential novelty effects during generalisation trials (given these words are not seen during encoding) needed to be reduced. Therefore, a SCT was implemented. Here, participants classified all 240 words used within the precision task as either human-made or natural before taking part in the precision task itself. Having participants be exposed to all the items used in the precision task should reduce item (though not associative) novelty effects and allow for firmer conclusions to be drawn about any activity differences in the hippocampus.

A new behavioural task that requires participants to categorise at the superordinate (human-made vs natural) level may influence memory and generalisation behaviour to differ from those observed in Chapter 2. During the behavioural work conducted in Chapter 2, only 4 (of 261 or 1.53%) participants explicitly mentioned the two superordinate categories: human-made and natural. Instead, most participants mentioned lower-level categories (e.g., animal, scenery, stationary). A concern was how participants using these superordinate categories more explicitly might change the previously observed behaviour. As such, the present pilot experiment was conducted to ensure the inclusion of the SCT did not alter the pattern of behaviour previously observed.

The pilot followed a similar structure to Experiment 2 in Chapter 2, but with the inclusion of the SCT and a 3-hour delay between Study and Test to account for possible logistical issues when booking the scanner for the fMRI investigation. Presuming the SCT task does not change memory or generalisation behaviour, the following behavioural effects should be observed: (1) greater memory accessibility but reduced memory precision in the clustered compared to the non-clustered condition (2) less divergence from the experimentally imposed cluster (i.e., von Mises distribution) for clustered novel items compared to the non-clustered novel items, and (3) non-clustered responses should show reduced kernel density at the centre of the experimentally imposed cluster indicating an avoidance behaviour.

4.3.2. Methods

4.3.2.1. Participants

Twenty-four participants (21 female) with a mean age of 19.42 years (*SD* = 0.76) took part. One participant did not return for the test phase, one was excluded due to an insufficient number of novel items being responded to, and the other due to their data not converging during the mixture model procedure. As such, the final sample consisted of 21 participants (19 female) with a mean age of 19.48 years (*SD* = 0.79). All participants were fluent English speakers with normal or corrected to normal vision and no known neurological condition. Participants took part in exchange for course credit or cash payment. Ethical approval was granted by the Department of Psychology Ethics Committee at the University of York.

4.3.2.2. Materials

The same 240 words from Chapter 2 were used. There were two semantic categories (120 words per category): human-made (e.g., chair, table) and natural (e.g., bee, mountain). Words were selected by assessing the semantic similarity (estimated via Euclidian distance) between vectors, obtained via a pre-trained word2vec model (Mikolov et al., 2013). Semantic distances within a category list were small, whilst the distance between lists was large. Sub-lists of 30 items were then generated, again controlling for semantic distance.

4.3.2.3. Procedure

4.3.2.3.1. Semantic Categorisation Task

Participants first completed the semantic categorisation task (see Figure 4.1). Participants were sequentially presented with all 240 words and asked to categorise each as either human-made or natural. The order of presentation was randomised for each participant. Each trial began with a fixation cross (1s) followed by the word to be categorised (< 2s). When the word was presented, participants had 2s to categorise the word using a keypress of 0 or 1. The key associated with a given category was counterbalanced across participants. Feedback was given to participants as to whether they provided a correct or incorrect categorisation on each trial; this was given via the word “Correct” or “Incorrect” appearing on screen for 0.5s in either green or red font, respectively. If no response was provided, the trial was deemed incorrect, feedback was given, and the subsequent trial began. Participants were given up to 1-minute breaks after every 60 trials, though they could skip these breaks if they wished to and continue with the task. Once completed, the Study Phase began. Participants were given 10 practise trials to complete prior to the SCT in order to familiarise them with the task; these trials used abstract nouns.

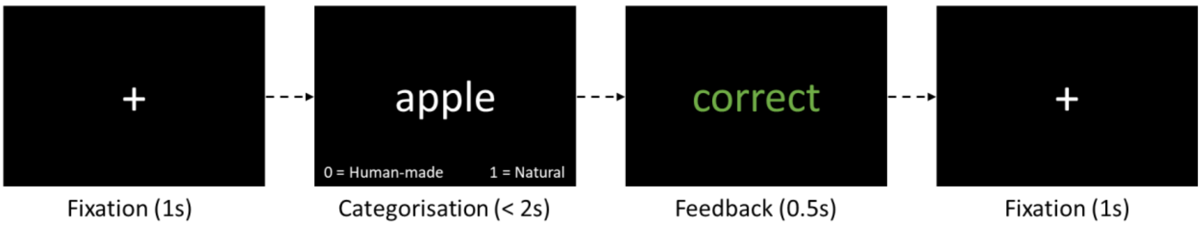

Figure 4.1. **Semantic Categorisation Task**. Demonstration of the trial structure for the semantic categorisation task. Participants would be presented with a fixation cross (1s) followed by a word that they needed to categorise (< 2s) and then feedback (0.5s) before moving on to the subsequent trial.

4.3.2.3.2. Precision Memory Task

4.3.2.3.2.1. Study Phase

The study phase was identical to that reported for Chapter 2: Experiment 2. Participants were asked to learn associations between words and locations around a circle. A total of 180 words were presented with one category of words belonging to the clustered and the other the non-clustered condition. The clustered condition had locations sampled from a von Mises distribution with a fixed width (κ = 2) and fixed, though participant-unique, mean. In the non-clustered condition, locations were sampled from a uniform distribution meaning there was no underlying pattern. The order of presentation was randomised across participants. During each trial, a fixation cross would appear (1s), followed by a location (2s), the word alone (4s) and then participants were asked to reposition a randomly located marker back to the location that had just been presented (< 6s). The trial was repeated if participants did not respond within the 6s time window or selected an area greater than 10° from the presented location. Practise trials took place before the Study Phase to ensure participants were familiar with the task. The practise words were abstract nouns to ensure they could not be confused with any study items. Participants were then instructed to return to the lab ~3 hours after completion to commence the Test Phase.

4.3.2.3.2.2. Test Phase

As the fMRI study would likely have some logistical constraints meaning participants could not complete the Test Phase immediately, the behavioural pilot implemented an approximate 3-hour gap between Study and Test. This gap was used to estimate the possible behavioural patterns present during the fMRI experiment. The average time between Study and Test was 2.87 hours (*SD* = 0.55).

Upon returning, participants were asked to recall the previously presented word-location associations (180 old words) and give a location to 60 novel words they had not previously associated with a location. Novel words still belonged to the same semantic groupings described above and were presented during the semantic categorisation task. Old and new words were intermixed with presentation order randomised. On each trial, participants were presented with a fixation cross (1s), the word alone (2s) and then asked to reposition a randomly located marker back to a remembered location or to make a best guess (< 10s). Participants were not told about the novel words at test, with the trial structure being identical.

4.3.2.3.3. Introspection Questionnaire

Finally, participants were asked to complete the Introspection Questionnaire; this was identical to Chapter 2 and probed participants' perceptions of the task (e.g., whether they perceived a pattern or the presence of novel items).

4.3.3. Data Handling

4.3.3.1. Mixture Model Estimation

Mixture modelling was applied to the memory data to estimate accessibility (i.e., retrieval probability) and precision (i.e., how precisely are locations remembered for retrieved items) for each participant in each condition, separately. The mixture model is described in Chapter 2. In brief, the algorithm used the replacement error for each response to estimate each metric using an expectation maximisation algorithm. Two models could be fit to the data to obtain the accessibility and precision metrics. If the first model provided a good fit, the parameter estimates were used. However, if the fit was not adequate, an alternative fitting procedure was used. If neither model provided a good fit, the dataset was excluded. Once both p and κ parameters were derived, they were converted to entropy measures: I_p and I_κ. These metrics were then used to compute an overall metric of performance: total information (I_t).

4.3.3.2. Kernel Density Estimation

Kernel density estimates were computed as in Chapter 2. These are nonparametric representations of the locations selected by participants. To compute these, individual von Mises probability density functions, with a concentration of κ = 2, were centred on each location. These acted as a smoothing kernel to get probability density estimates at any angle around the circle and were used to plot the distribution of locations centred to the imposed cluster and estimate the Kullback-Leibler diverge values (see below).

4.3.3.3. Kullback-Leibler Divergence

The kernel density estimates were used to assess the similarity between the distribution of locations of generalisation trials and the experimentally imposed von Mises distribution using Kullback-Leibler Divergence (D_{KL}). D_{KL} measures divergence between two distributions, with higher values representing greater divergence (i.e., less similarity) between distributions.

4.3.3.4. Exclusion Criteria

The same exclusion criteria from Chapter 2 were used. Briefly, individual trials from a dataset were excluded when participants did not provide a response during that trial or when the trial had been repeated five times or more during study. Datasets would only be included for analysis when the following criteria were met: (1) both the study and test trials were complete, (2) the memory trials (for clustered and non-clustered separately) did not have fewer than 70 responses, (3) the generalisation trials (for clustered and non-clustered separately) did not have fewer than 15 responses, (4) the dataset was not corrupted, and (5) the mixture model could be fit adequately to the data.

4.3.4. Statistical Analysis

GLME models were used to analyse the data. To assess whether changes in behaviour were present, the data from Chapter 2 was input into all models. Six models were generated, four of which used the same fixed effect and random effect structure, with the other two using an alternative fixed and random effect structure. All models were fit to the data using a log link function, a gamma distribution to model data spread and estimated using the maximum likelihood fitting method within the MATLAB Statistics and Machine Learning Toolbox.

The first four models analysed: (1) Total Information, (2) Accessibility, (3) Precision, and (4) D_{KL} von Mises for novel items. The models included three dummy-coded fixed effects:

Clustering (0 = Non-Clustered; 1 = Clustered), Delay (with 0hrs being the reference variable and 3-hrs and 24-hrs being explicitly modelled) and Setting (0 = In-Lab, 1 = Online). There was no 3-hr online fixed effect, so a design matrix of full rank could not be generated. Therefore, clustering was made to interact with delay and setting, but delay and setting did not interact with each other. In addition to these fixed effects, two random effects were estimated. One random effect allowed the intercepts to vary based on the participant, and the other allowed the slopes to vary based on the effect of clustering. All elements of the associated random effects covariance matrix were estimated via the data. As the variables included in the model were dummy coded, linear hypothesis tests of the model coefficients were undertaken to examine the effects of interest. The contrast matrices used for these comparisons are detailed in Appendix A.

For the remaining two models, which analysed the kernel density of non-clustered items at the cluster centre, the effects of delay and setting were included, with random intercepts based on each participant. For this analysis, differences based on delay and setting were analysed, along with identifying whether kernel density differed from that expected of a circular uniform ($2\pi^{-1}$). No interactions were modelled. Again, the contrast matrices used can be found in Appendix A.

For all analyses, the simple effect of clustering isolated to the pilot data was first explored; this was done to assess the effect of clustering independent of data from Chapter 2. For the kernel density (i.e., avoidance) effects, the difference from the uniform value isolated to the pilot data was first analysed. Subsequently, any change between mean estimates from the behavioural pilot and the data from Chapter 2 were investigated; this was done by collapsing 0-hrs and 24-hrs into one level (i.e., Chapter 2) and comparing this to the pilot data (3-hr delay). Finally, an assessment of whether the effect of clustering had changed across chapters was undertaken by looking at the interaction between clustering and delay (collapsed as previously specified) where available. Though other effects were present in the model (e.g., setting), they were not of interest to the present analysis. However, they were included to ensure the model was correctly specified to derive unbiased parameter estimates.

All mean values represent the estimated marginal means derived from the GLME. Further, the Cohen's *d* and Bayes Factors reported were calculated using the method described in Berens et al. (2020) and estimated only on the fixed factors. The Bayes Factors used the default parameters: a prior Cauchy distribution *r* = .707, centred at 0. All analyses used two-tailed tests unless otherwise specified.

4.3.5. Results

4.3.5.1. Memory

An analysis of the memory measures (I_t, I_p, and I_k) was undertaken first. Figure 4.2 below shows the pattern of results for the pilot compared to the data in Chapter 2.

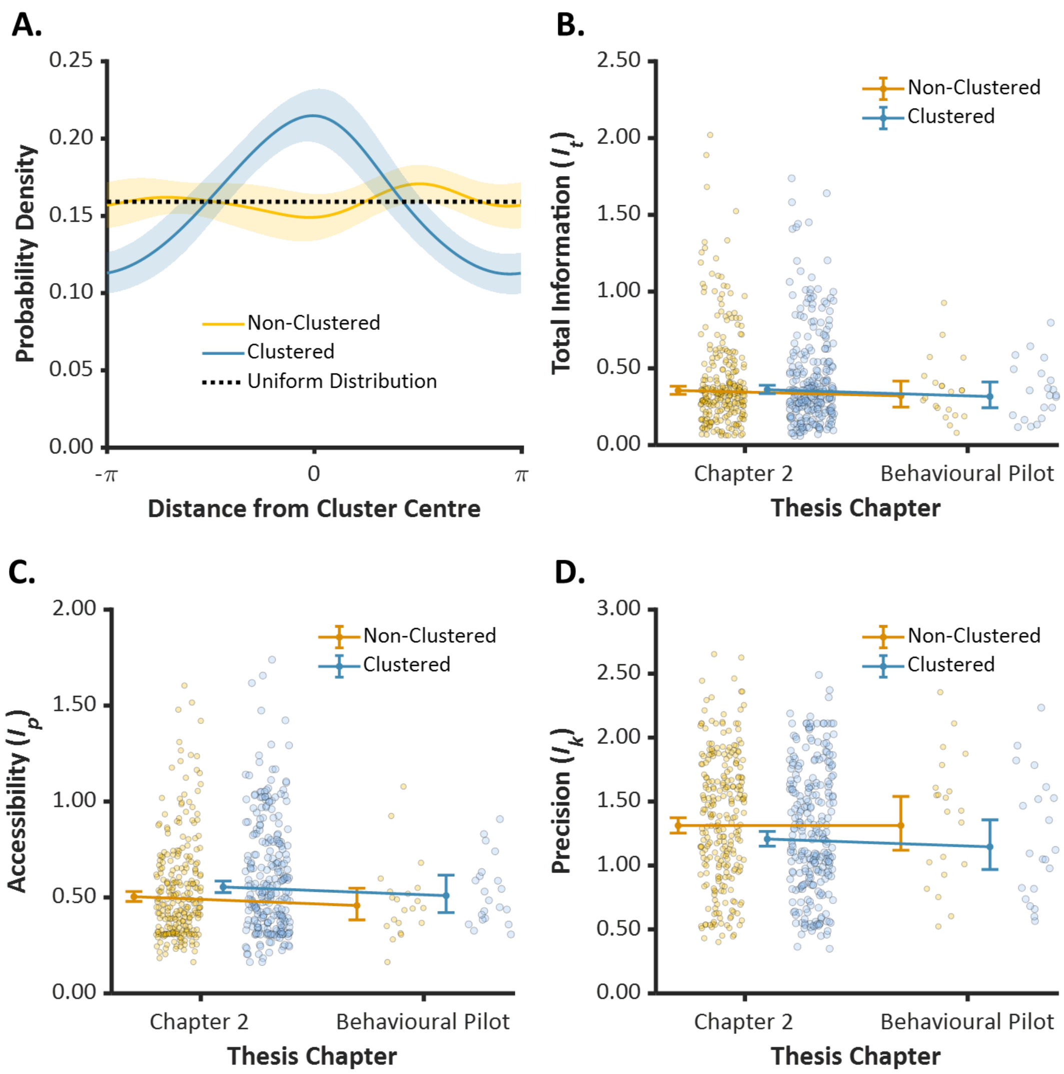

Figure 4.2. **Assessment of memory performance across chapters**. (A) **Kernel Density Plot**: This represents the average spatial distribution of locations selected for old (memory) items, centred to the experimentally imposed von Mises distribution for the pilot data alone. (B-D) **Memory Metrics**: Illustrating the effects of clustering and chapter on: (B) Total Information, (C) Accessibility, and (D) Precision. The chapter variable was computed by averaging 0-hrs and 24-hrs (Chapter 2) into one level and comparing this to 3-hrs delay (Behavioural Pilot). Error bars represent the estimated marginal means and 95% confidence intervals of the model. Dots represent individual data points from a given participant.

First, an exploration of total information occurred. Examination of the effect of clustering for the pilot only found no significant difference between conditions, $t(556) = 0.13$, $p = .898$, $d = 0.02$, $BF_{01} = 6.53$. This effect is in line with the results of Chapter 2, whereby both clustered and non-clustered conditions showed similar levels of total information. Next, the effect of delay was examined, finding no significant difference between estimates of total information, $t(556) = 0.92$, $p = .339$, $d = 0.11$, $BF_{01} = 4.98$. This suggests that the levels of total information in the behavioural pilot are consistent with those in Chapter 2. Finally, no significant interaction between clustering and delay was found, $t(556) = 0.25$, $p = .806$, $d = 0.04$, $BF_{01} = 5.25$. These results suggest that the SCT inclusion has not affected the pattern of results related to total information.

Next, accessibility was assessed. Here, no significant difference between clustered and non-clustered conditions were found when analysing only the pilot data, $t(556) = 1.22$, $p = .225$, $d = 0.17$, $BF_{01} = 3.30$. However, the mean estimates for the clustered and non-clustered conditions were in the expected direction (see Figure 4.2). Support for a lack of change comes from the lack of significant difference between Chapter 2 or the pilot data, $t(556) = 1.05$, $p = .296$, $d = 0.13$, $BF_{01} = 4.50$. Further, no significant interaction between clustering and delay was observed, $t(556) = 0.12$, $p = .902$, $d = 0.02$, $BF_{01} = 5.36$. These results suggest that the effect of clustering had not changed across chapters. Like total information, it is apparent that the inclusion of the SCT has not impacted levels of accessibility.

Finally, memory precision was analysed. Assessing the effect of clustering on only the pilot found no significant difference between conditions, $t(556) = 1.22$, $p = .223$, $d = 0.17$, $BF_{01} = 3.28$. However, like accessibility, the direction of effects was as expected (see Figure 4.2) and in line with Chapter 2. This was supported by a lack of difference in mean estimates between Chapter 2 and the pilot data, $t(556) = 0.40$, $p = .693$, $d = 0.04$, $BF_{01} = 8.26$. Additionally, the effect of clustering across chapters did not show a significant change, $t(556) = 0.45$, $p = .652$, $d = 0.08$, $BF_{01} = 4.92$. These effects support the view that the inclusion of the SCT did not influence any of the memory measures used within the current paradigm.

4.3.5.2. Generalisation

To examine generalisation, the locations selected for novel words across the two conditions were compared to the von Mises distribution to assess which was more similar to (i.e., less divergent from) this distribution. It was expected that the clustered condition would show less divergence than the non-clustered. However, examination of the pilot data found no such difference, $t(556) = 1.18$, $p = .239$, $d = 0.16$, $BF_{01} = 3.44$. This result suggests that both conditions were equally divergent to the von Mises distribution. However, as with the memory analyses, the direction of effect was as expected (see Figure 4.3), implying the SCT has not influenced generalisation behaviour. To corroborate this conclusion, the D_{KL} values obtained in the present pilot were compared to the data from Chapter 2, finding no significant difference, $t(556) = 0.51$, $p = .609$, $d = 0.05$, $BF_{01} = 7.87$. Additionally, no interaction between clustering and delay was found, $t(556) = 1.12$, $p = .264$, $d = 0.19$, $BF_{01} = 3.06$. These results suggest that the effect of clustering had not changed across chapters. Therefore, like memory, the inclusion of the SCT did not have a clear effect on generalisation behaviour in the present pilot experiment.

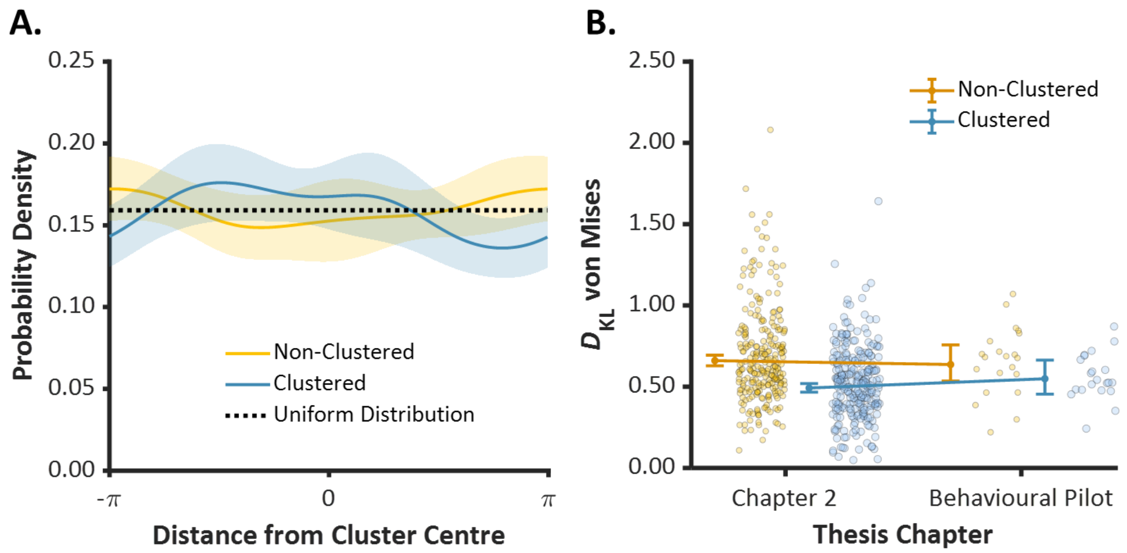

Figure 4.3. **Assessment of generalisation performance across chapters**. (A). **Kernel Density Plot**: The average spatial distribution of locations selected for generalised (novel) items centred to the experimentally imposed von Mises distribution for the pilot data only. (B) **D_{KL} von Mises**: Estimated marginal means from the GLME as a function of chapter and clustering. The chapter variable was computed by averaging 0-hrs and 24-hrs (Chapter 2) into one level and comparing this to 3-hrs delay (Behavioural Pilot). Error bars represent the estimated means and 95% confidence intervals of the model. Dots represent individual data points from a given participant.

4.3.5.3. Avoidance

Next, evidence of avoidance was assessed by comparing the probability densities of non-clustered locations at the centre of the cluster to that of a uniform density. The distribution of locations selected for memory trials in the pilot data are shown in Figure 4.2 (above). Examination of the selected locations for the pilot data found no significant evidence of avoidance, $t(278) = 1.38$, $p = .084$ (one-tailed), $d = 0.15$, $BF_{01} = 3.35$. Despite the Bayes Factor suggesting moderate support for the null, the mean estimate is in the expected direction (see Figure 4.4). Further, no significant differences between estimates from the pilot and Chapter 2 were observed, $t(278) = 0.57$, $p = .567$, $d = 0.11$, $BF_{01} = 4.34$. Overall, these findings suggest little evidence that the SCT impacted avoidance behaviour. Though the Bayes Factor provided moderate evidence, the lack of a delay effect supports the view that there is little to no change across chapters.

For generalisation, the locations selected for both clustered and non-clustered conditions are shown in Figure 4.3 (above). Similar to memory, there was no significant evidence of avoidance when isolated to the pilot data, $t(278) = 0.99$, $p = .162$ (one-tailed), $d = 0.11$, $BF_{01} = 5.24$. Again, the Bayes Factor provides moderate evidence favouring a null effect, contrary to expectations given the previous work. However, examination of across chapter differences found no significant effects, $t(278) = 0.52$, $p = .601$, $d = 0.10$, $BF_{01} = 4.44$. The lack of difference in kernel density values across chapters and examination of the mean estimates (Figure 4.4) suggest the data may have been too underpowered to find any significant effect.

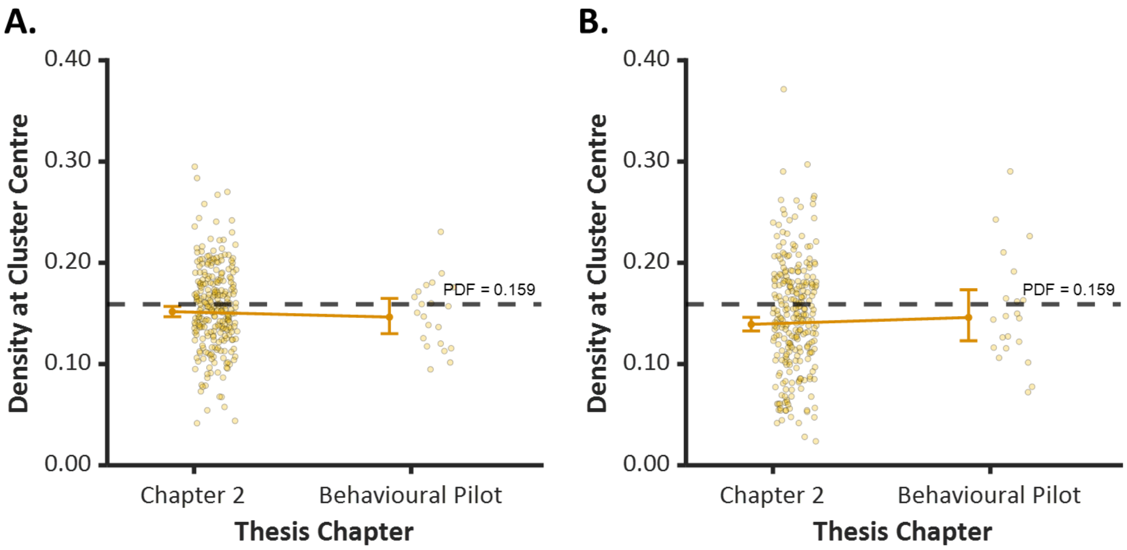

Figure 4.4. **Assessment of avoidance behaviour across chapters**. (A) **Memory**: This represents the mean density for chapter 2 and the present behavioural pilot for memory trials. (B) **Generalisation**: This represents the mean density for both chapters for generalisation trials. As before, Chapter 2 was the combination of 0-hrs and 24-hrs delay. Error bars represent the estimated marginal means and 95% confidence intervals. Dots represent the individual participant scores.

Overall, these results support the view that the inclusion of the SCT has not impacted the behaviour observed in previous experiments (i.e., Chapter 2). The moderate evidence for a null effect in some of the analyses may be worrisome. However, as there was no evidence of differences across chapters, the likely reason for these moderate Bayes Factors could be the product of noise in the pilot sample.

4.3.6. Discussion

The behavioural pilot aimed to investigate whether implementing the SCT would impact the behaviour reported previously (see Chapter 2). The results show no evidence of any meaningful behaviour change compared to the data collected for Chapter 2, suggesting the inclusion of the SCT had little impact on the behaviours of interest. Numerically, the clustered condition showed greater accessibility but reduced precision compared to the non-clustered condition. Further, when generalising, participants showed greater adherence to the von Mises distribution in the clustered compared to the non-clustered condition. Finally, there was evidence of avoidance behaviour, both in memory and generalisation. Though the moderate Bayes Factors in favour of the null may appear worrisome, the lack of differences between the present pilot and the data reported in Chapter 2 means no firm conclusions can be made without further data.

One of the concerns with introducing the SCT was the possibility that participants' explicit use of the superordinate groupings (i.e., human-made and natural) may influence some of the behavioural patterns found in Chapter 2. In Chapter 2, very few participants, 4 out of 261, acknowledged the presence of these superordinate categories, instead focusing on the more basic categories (e.g., fruit, office items). In contrast, 29% (6/21) of participants reported using these superordinate categories during the pilot debrief when dealing with inferences for the task. Evidently, explicit acknowledgement of the superordinate categories may not drastically change the behaviour observed in the precision task. Of course, caution should be taken with this interpretation given the small sample size.

Overall, the behavioural pilot suggests that the implementation of the SCT may not impact the behaviour observed in the precision paradigm. Therefore, the SCT may be a valuable mechanism for reducing some of the novelty effects associated with the generalisation trials and allow us to make firmer conclusions as to why the hippocampus, if active, may have been recruited for those trials (Strange et al., 2005).

4.4. fMRI Pilot

4.4.1. Aims and Hypotheses

Due to the COVID-19 pandemic, the fMRI experiment moved to an exploratory investigation of pilot data as a means of generating testable hypotheses for an independent experiment. The present study used a Clustering (Non-Clustered vs Clustered) x Trial Type (Memory vs Generalisation) design to investigate memory-based generalisation. Implementing this design would allow us to investigate any neural dissociations between memory and generalisation as a function of whether items belonged to the clustered or non-clustered conditions. Despite no firm hypotheses for the present exploration, a Region of Interest (ROI) analysis was conducted on the vmPFC and hippocampus, given their involvement in the literature previously discussed (see Introduction).

4.4.2. Methods

4.4.2.1. Participants

A total of 5 participants (4 female) were recruited. The mean age of the sample was 24.80 years (*SD* = 3.97). One participant's behavioural data did not converge during mixture modelling and was removed for any behavioural analysis reported. However, given the limited sample, the participant was included in the fMRI analysis. Therefore, for behavioural analyses, there was a total of 4 participants (3 female) with a mean age of 26.00 years (*SD* = 3.54). All participants were fluent English speakers with normal or corrected to normal vision and no known neurological condition. Participants took part in exchange for cash payment. Ethical approval was granted by the Research Governance Committee in the York Neuroimaging Centre.

4.4.2.2. Procedure

Upon arrival, participants completed the SCT and Study Phase of the precision task, identical to the description above. Participants were asked to return ~3 hours later to complete the Test phase within the scanner. The average time between Study and Test was 2.67 hours (*SD* = 0.83).

The Test Phase generally followed the same structure as described previously. However, the trial timings were adjusted; these changes are illustrated in Figure 4.5 below. These timings were similar to the online experiments described in Chapter 2. The inter-trial interval (ITI) was fixed to 7.5s; this consisted of a 0.5s fixation cross, the word alone (1s) and then the circle appearing around the word with a location marker placed randomly around its circumference, it was here participants could make a response (< 6s). During the response window, participants moved the marker via keypresses to either the location they remembered or made a best guess, indicating a final position via another keypress. Once a response was given, the circle and marker would dim, and the subsequent trial would begin once the full ITI had elapsed (i.e., 6s had passed since being able to make a response). If a response were not given within the time window, the subsequent trial would begin after the ITI had elapsed.

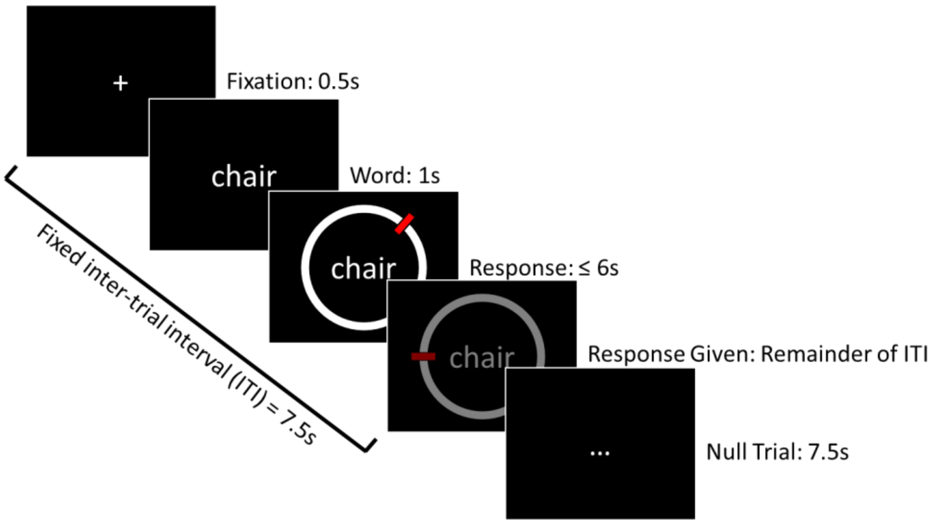

Figure 4.5. **Test Phase: fMRI Experiment**. Schematic of the structure for test trials during scanning. Trial length was fixed to 7.5s, with experimental trials starting with a fixation cross (0.5s), the word alone (1s) and then providing participants with the opportunity to reposition the marker back to the location they remember or to make a best guess (< 6s). Once a response was given, the circle, marker and word would change in luminance to signify a response was registered, and the trial would not continue until the ITI had elapsed. For null trials, an ellipse (…) appeared on the screen for the full ITI.

There were two functional runs. Each run consisted of 132 trials: 90 memory and 30 generalisation, split evenly across levels of clustering. There were also 12 null trials per run. During a null trial, participants would be presented with an ellipse (…) on screen for the full ITI. The timing and order of trials within each run was determined using optseq2 (Dale, 1999; http://surfer.nmr.mgh.harvard.edu/optseq/) to optimise estimation efficiency. Two trial sequences were developed via optseq2. Participants experienced both sequences, but the order was counterbalanced across participants. The items presented during each run was still randomised.

4.4.2.3. MRI Acquisition

All functional and structural volumes were acquired on a 3T Siemens MAGNETOM Prisma scanner equipped with a 32-channel phase array head coil at the York Neuroimaging Centre at the University of York. Functional data were acquired using T2*-weighted echo-planer imaging (EPI) via GRAPPA parallel imaging with an acceleration factor of 2 and a multi-band acceleration factor of 2. A total of 48 axial slices (0° tilt from AC-PC line) per volume were acquired in an interleaved order with the following parameters: TE: 26ms, TR: 1200ms, Flip Angle: 75°, Field of View (FOV) = 192 x 192mm, Slice Thickness: 3mm, Acquisition Matrix: 64 x 64. Two functional runs (825 scans lasting 16.5 minutes each) were collected per participant.

To unwarp the data, gradient echo field maps were acquired for each participant with the following parameters: TE1 = 4.92ms, TE2 = 7.38ms, TR = 650ms, Flip Angle = 60°, FOV = 192 x 192mm, 48 slices with a slice thickness of 3mm. Additionally, for the purposes of co-registration and normalisation, one high-resolution T1 weighted structural image was acquired using a magnetisation prepared rapid gradient echo (MP-RAGE) pulse sequence with the following parameters: TR = 2300ms, TE = 2.26ms, Flip Angle = 8°, FOV = 256 x 256mm, Image Resolution = $1mm^3$.

4.4.3. Data Handling and Statistical Analysis

4.4.3.1. Behavioural Analysis

The same behavioural analyses as reported in the behavioural pilot above were conducted. In short, a GLME was applied to the data using a gamma distribution to model the spread of the data. A log-link function was applied to define the relationship between the dependent variable and the linear combination among variables. Although previous models have included two random effects per participant, this led to overfitting in the present data. Therefore, all models applied a random intercept for each participant, with no random slopes.

4.4.3.2. fMRI Pre-Processing

Image pre-processing was performed using the Statistical Parametric Mapping (SPM12; Wellcome Department of Cognitive Neurology, London, United Kingdom) via MATLAB (2019b; Mathworks Inc., Natick, MA, United States). First, EPIs were bias-corrected by segmenting the first image and applying the bias field to all subsequent EPI's. Following this, EPIs were corrected for head motion (realignment), aligned to the first scan of each run, and corrected for magnetic field inhomogeneities (unwarping) using voxel-displacement maps derived from the phase and magnitude field maps of each subject. The structural T1 image was then co-registered to the mean EPI. A manual reorientation of the images was then undertaken to ensure volumes were oriented to the anterior commissure. EPI images were then spatially normalised to Montreal Neurological Institute (MNI) space with transformation parameters derived from warping each participant's structural image to a T1-weighted average template image (using the DARTEL toolbox; Ashburner, 2007). Normalised images were spatially smoothed using an isotropic Gaussian kernel of 8mm full width at half maximum.

4.4.3.3. fMRI First-Level Analysis

Subject-specific models were constructed where trials were modelled by convolving a boxcar function, based on stimulus onset and event duration, with a canonical hemodynamic response function. Though trial lengths were fixed to 7.5s, the duration time used was from stimulus onset (i.e., when the word appeared on-screen) to participant response. In instances where a response was not given, the entire 7s interval was modelled, ignoring the fixation cross. The model contained four regressors of interest for each functional run representing the variables: non-clustered memory, clustered memory, non-clustered generalisation and clustered generalisation. Six additional regressors were also included representing movement parameters estimated during spatial alignment (three rigid-body translations and three rotations). Voxel-wise parameter estimates for each regressor were obtained via restricted maximum-likelihood estimation, using a temporal high-pass filter (128Hz) to remove low-frequency drifts and applying a first-order autoregressive model (Friston et al., 2002) to account for temporal autocorrelation.

4.4.3.4. MRI Group-Based Analysis

4.4.3.4.1. Whole-Brain

A whole-brain voxel-wise analysis was conducted. First-level contrasts of the parameter estimates for each of the four experimental conditions (averaged across the two functional runs) for each participant were entered into a second-level clustering (non-clustered and clustered) and trial type (memory and generalisation) repeated measures ANOVA to examine both main effects and the interaction. All effects were thresholded at $p < .001$ uncorrected, with an extent threshold of 10 voxels.

4.4.3.4.2. ROI Definition

Given previous work implicating both the hippocampus and vmPFC in schema use, anatomical masks were created using the automated anatomical atlas (AAL; Tzourio-Mazoyer et al., 2002) implemented in the Wake Forest University (WFU) PickAtlas Toolbox (Maldjian et al., 2003). Both the left and right hippocampus were included for the hippocampal ROI. For the vmPFC, as there is no devoted label, both left and right gyrus rectus and the left and right medio-orbital section of the frontal cortex were used; this is identical to other experiments (e.g., Liu et al., 2016; Raykov et al., 2020). The mean beta values for each experimental condition were then extracted using MarsBaR (Brett et al., 2002). A linear mixed model (LMM) was applied to these estimates to examine the effects of interest, with random intercepts for subject. The LMM used a restricted maximum likelihood estimation fitting method, with the covariance pattern estimated using the log-Cholesky parameterisation; this was similar to the covariance pattern used for the GLME's. Estimation of the LMM used the MATLAB Statistics and Machine Learning Toolbox. Cohen's *d* and Bayes Factors were computed similarly to the GLME's using only the fixed effects. Note, given the limited sample, caution should be taken with interpreting all effects reported.

4.4.4. Results

4.4.4.1. Behavioural Analysis

4.4.4.1.1. Memory

First, the behavioural effects for memory were assessed. For total information, the clustered (M = 0.62, SE = 0.27) was numerically greater than the non-clustered (M = 0.47, SE = 0.27) condition, but not significantly different, $t(6) = 1.61$, $p = .158$, $d = 1.14$, $BF_{01} = 0.92$. The same pattern of effects was found for accessibility, with clustered (M = 0.76, SE = 0.29) being numerically larger than non-clustered (M = 0.58, SE = 0.29), $t(6) = 1.66$, $p = .148$, $d = 1.17$, BF_{01} = 0.88. Both results are in line with the pattern of effects previously observed. However, for precision, the clustered condition (M = 1.52, SE = 0.07) showed greater numerical precision compared to the non-clustered (M = 1.50, SE = 0.07), $t(6) = 0.17$, $p = .873$, $d = 0.12$, $BF_{01} = 1.90$; this is contrary to the pattern typically found. Nevertheless, despite the mean estimates suggesting a different direction of effects for precision, the difference of .02 in the mean estimate suggests this may be the product of noise due to the small sample (N = 4 for the behavioural analyses).

4.4.4.1.2. Generalisation

Examination of locations selected for novel items found that clustered items (M = 0.41, SE = 0.19) were placed more similarly to the experimentally imposed pattern (i.e., von Mises distribution) than the non-clustered (M = 1.24, SE = 0.19) items, $t(6) = 4.05$, $p = .007$, $d = 2.86$, $BF_{01} = 0.005$. This finding is in line with previous work reported throughout the thesis and shows that participants may be able to generalise the pattern presented at study to novel items.

4.4.4.1.3. Avoidance

Finally, an examination of whether the probability densities of non-clustered trials, centred to the experimental cluster, differed from a uniform density in both memory and generalisation was undertaken. For memory trials, there was significantly reduced density for non-clustered items (M = 0.11, SE = 0.08), $t(3) = 4.14$, $p = .013$ (one-tailed), $d = 2.07$, BF_{01} = 0.003. This was also true for generalisation trials (M = 0.06, SE = 0.24), $t(3) = 3.89$, $p = .015$ (one-tailed), $d = 1.95$, $BF_{01} = 0.007$. These results are both in line with expectations, showing evidence of avoidance in placing non-clustered items close to the cluster centre.

4.4.4.2. fMRI Analysis

Table 4.1 shows the output for the whole-brain voxel-wise analysis. For the present analysis, all peak-level activations that were significant at an uncorrected $p < .001$ threshold are presented. However, the p-values reported are corrected for family wise-error. Region labels were acquired using the AAL. However, the region labelled as putamen via the AAL also contained the caudate nucleus (see Figure 4.6). As such, this area of activation is referred to as the dorsal striatum to encapsulate both regions. From the whole-brain analysis, clustering had no suprathreshold effects. However, both trial type and the interaction with clustering did. Areas such as the dorsal striatum, ventral striatum and cerebellum were significantly active before correction. However, following family-wise error correction, only the dorsal striatum showed marginal significance (p = .057).

Table 4.1.

Statistical Results of the Whole Brain Analysis, with labels from the AAL Atlas.

Region	Voxels	MNI Coordinates			z	p_{FWE}
		X	Y	Z		
Type						
L Putamen	464	-18	18	-4	4.88	.057
R Caudate	87	22	28	8	4.26	.394
R Ventral Striatum	120	14	12	-8	3.98	.689
L Cerebellum	13	-16	-56	-42	3.34	.998
Clustering						
No suprathreshold effects present						
Type x Clustering						
R Putamen	63	-10	24	0	3.71	.914
L Caudate	11	0	10	8	3.45	.992

Note. Effects thresholded at p < .001 uncorrected and an extent threshold of 10 voxels. Voxels are listed in MNI standard space. L = Left, R = Right.

MarsBaR was used to extract the beta parameters for the dorsal striatum using the peak voxel coordinates (reported in Table 4.1, above) with a 10mm sphere. The beta parameters were entered into a linear mixed model as described for the anatomical ROI's (see ROI Definition). As the betas were selected from the peak voxel coordinates from the trial type contrast using the repeated measures ANOVA, this effect was not reported. However, the direction of the effect was assessed, which showed that memory trials had greater activation than generalisation (see Figure 4.6). Nevertheless, no main effect of clustering (F(1,16) = 3.65, p = .074, d = 0.62, BF_{01} = 0.75), nor an interaction (F(1,16) = 3.35, p = .086, d = 0.95, BF_{01} = 0.75) were found. However, these were marginally significant.

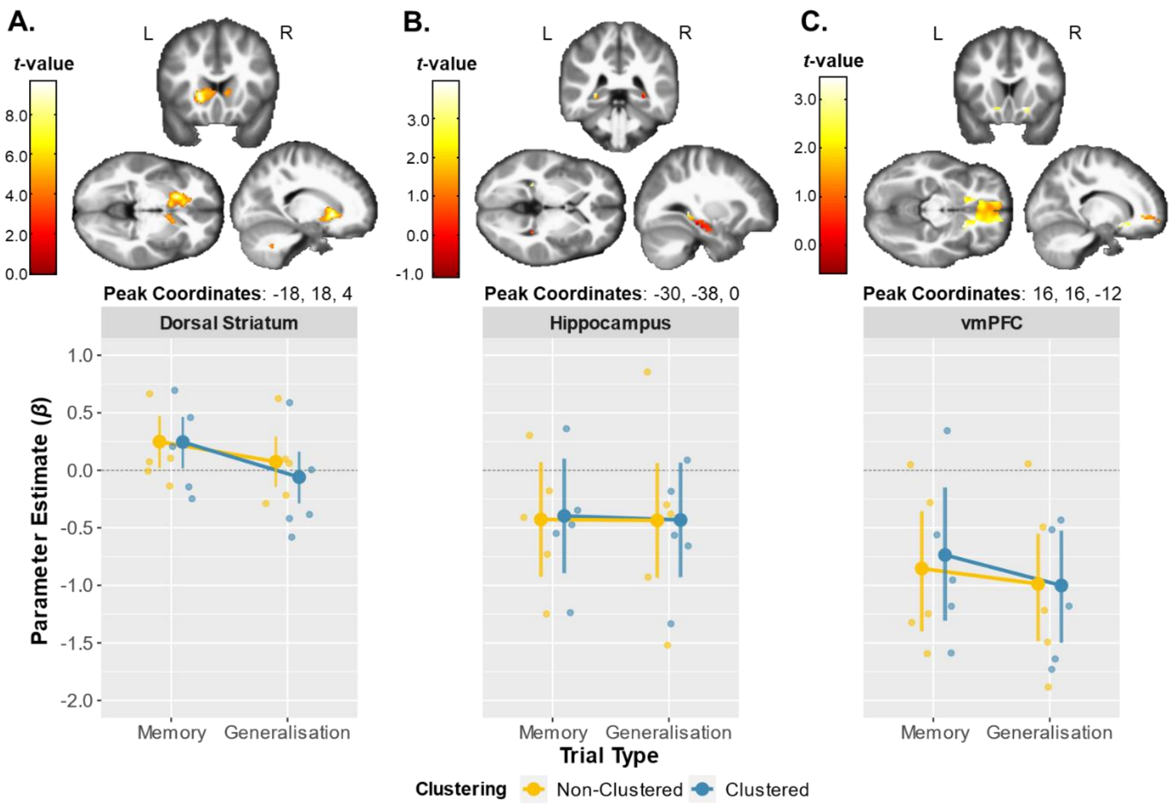

Figure 4.6. **Whole-brain and ROI Analysis Heatmaps and Descriptives**. Above each of the error-bar plots are the peak activations within each region: (A) Dorsal Striatum, (B) Hippocampus, and (C) vmPFC viewed from the axial, sagittal and coronal planes. Heatmaps represent the *t*-test contrast for memory trials greater than generalisation trials. The error-bar plots show the mean estimates derived for each ROI as a function of clustering and trial type. *t*-value colour bars are presented to show the extent of activation within an ROI. Across all plots, dots represent the beta parameters for individual subjects. Error bars represent the 95% CI around the estimated mean.

Subsequently, the anatomical ROIs for the hippocampus and vmPFC were analysed to examine whether any effect of clustering, trial type or interaction were present. Figure 4.6 shows the mean activation for these comparisons.

For the hippocampus, there was no significant difference in BOLD activity as a function of clustering, ($F(1,16) = 0.03$, $p = .870$, $d = 0.05$, $BF_{01} = 3.14$), trial type ($F(1,16) = 0.04$, $p = .840$, $d = 0.07$, $BF_{01} = 3.12$) nor an interaction between the two ($F(1,16) = 0.02$, $p = .900$, $d = 0.07$, $BF_{01} = 2.28$). These results suggest the hippocampus did not show differences in BOLD activation across conditions. In contrast, for the vmPFC, there was a main effect of trial type, with memory trials being associated with greater BOLD response than generalisation trials,

$F(1,16) = 5.49$, $p = .032$, $d = 0.76$, $BF_{01} = 0.35$. However, there was no main effect of clustering ($F(1,16) = 0.37$, $p = .549$, $d = 0.20$, $BF_{01} = 2.74$), nor an interaction ($F(1,16) = 0.60$, $p = .449$, $d = 0.40$, $BF_{01} = 1.89$). These patterns of results suggest the activation of the vmPFC during memory trials was not differentially affected based on whether the trial was clustered or non-clustered.

4.4.5. Discussion

Due to the COVID-19 pandemic, a preliminary analysis was undertaken on the existing fMRI data. Though firm conclusions cannot be drawn with this sample ($N = 5$), the results indicate possible differences in vmPFC and dorsal striatal activation based on trial type. In both instances, there was greater BOLD activation during memory trials compared to generalisation. However, no significant effects were found when assessing the hippocampal ROI, suggesting this region was equally as active (or inactive) across trials irrespective of condition. Below a brief discussion of possible explanations for these patterns of activation is provided.

Finding the vmPFC to be more active during memory than generalisation may relate to its involvement in interference (Incisa della Rocchetta & Milner, 1993; Spalding et al., 2018). In their study, Spalding et al., (2018) had patients with vmPFC lesions and controls complete an associative inference task. Participants learned associations across items: A-B and B-C. It was found that patients were impaired in inferring that A and C were related through B but unimpaired in their ability to recall associated pairs (e.g., B-C). However, despite memory appearing similar to controls in the task, patients did struggle to retrieve initially learned associative pairs (e.g., A-B) compared to retrieving the second pair (e.g., B-C). Due to the vmPFC damage, retroactive interference may have taken place, meaning that recall of the original A-B pairing was hindered compared to the more recently encoded item (i.e., B-C). This finding would suggest that the vmPFC serves a role in preventing interference among competing items and is in line with other work. Concerning the present work, this may explain why there was greater vmPFC activation during memory compared to generalisation trials. During memory, related information may become active and need to be suppressed so that location judgements are not biased by these overlapping events, but instead rely mainly on the individual experience. In contrast, during generalisation, allowing for more interference from several experiences may allow for greater generalisation, meaning vmPFC activation was decreased.

There was no evidence of differential hippocampal activation within the present paradigm. Based on the previous literature, the hippocampus was expected to show differences in activity, being more active during memory than generalisation (Aggleton & Brown, 1999; Korkki et al., 2021; Moscovitch et al., 2016; Richter et al., 2016). Finding no differential activation within the hippocampus may be due to signal dropout and lack of power (Olman et al., 2009) and/or due to not analysing the hippocampus along the anterior-posterior axes separately (Bowman & Zeithamova, 2018, 2020; Collin et al., 2015; Frank et al., 2019; Schlichting et al., 2015). This proposal is discussed in more length in the Lessons Learned section below.

Though only marginally significant, the dorsal striatum was more active during memory than generalisation; this may be related to its involvement in memory gating (McNab & Klingberg, 2008). In their network model, O'Reilly and Frank (2006) proposed that communication between the striatum and prefrontal cortex (PFC) ensures that items with high utility (i.e., leading to the correct response) are maintained in working memory when making a decision, while irrelevant information is not. The dorsal striatum and vmPFC are more active during memory than generalisation in the present work. Speculatively, there may be functional coupling between the two regions as the striatum attempts to modulate what items will be relevant in the current context and which are not, whilst the vmPFC modulates this process, either to decrease interference or support schematic influences on memory. Working together, the regions may allow for competing information to not bias location judgements during memory trials. However, during generalisation, allowing more information to enter working memory to help weight the location judgment may be beneficial to make a more accurate judgement.

Interestingly, no differences based on the effect of clustering – either alone or interacting with trial type – were found. One interpretation may be that in both the clustered and non-clustered conditions, participants perceived a pattern. This is evidenced by the avoidance effect within the non-clustered condition. The presence of the avoidance suggests the possibility that two patterns are perceived by participants – one in the clustered condition and the other in the non-clustered condition. Consequently, the two conditions are no longer different. Instead, the clustering manipulation will work similarly to other studies where two schemas on opposite sides of the circle were present (e.g., Tompary et al., 2020). As such, for the present work, finding a lack of differences based on clustering may be explained by the presence of the avoidance effect alone. Participants perceiving both conditions as having a pattern may result in both relying on the same underlying neural processes.

Though firm conclusions cannot be made, the preliminary findings of the present work suggest that the vmPFC and dorsal striatum may play important roles in the memory behaviour observed in our task. Based on previous work, the vmPFC may attempt to prevent interference, whilst the dorsal striatum gates working memory from irrelevant information. In this way, the two regions may serve complementary roles. Caution is needed though given this is a reverse inference based on previous research. Though there is little evidence of differences in hippocampal activation based on the conditions of interest, this may be due to how the region was analysed in the present case. Overall, these findings give preliminary insight into the neural correlates of memory-based generalisation using the precision task.

4.5. Lessons Learned

This chapter aimed to provide a preliminary investigation into the neural correlates of memory-based generalisation using the precision paradigm. This section examines the lessons learned from these preliminary analyses, including possible design changes, analyses plans, and open questions, which could be used in an independent experiment.

4.5.1. Summary of Findings

Two studies are reported in this chapter: the behavioural pilot and the fMRI pilot. The behavioural pilot was undertaken to investigate how the inclusion of the SCT may influence the behaviour observed within the precision paradigm. The SCT was included to deal with item novelty effects that may have explained hippocampal activation during our task (Strange et al., 2005). The pilot found no evidence of changes in behaviour resulting from the SCT being included, with the results from the pilot not being significantly different from those in Chapter 2. Given that only five fMRI datasets were collected, a preliminary analysis of the collected data was undertaken. Here greater activation in the vmPFC (significant) and dorsal striatum (marginally significant) during memory trials compared to generalisation. Surprisingly, there was no differential activation of the hippocampus during the task.

4.5.2. Design Considerations

During the previous discussion, it was speculated that the vmPFC and dorsal striatum might work together, given that they were both more active during memory trials. In their neural network model, O’Reilly and Frank (2006) proposed that the dorsal striatum and PFC work together to maintain relevant information and keep irrelevant information out of working memory. Speculatively, memory trials would require greater vmPFC and dorsal striatal activation to ensure less bias in the reported location where more weight should be given to the individual memory. In contrast, generalisation trials may require more information to enter working memory to make a location judgment. Therefore, in future, it may be helpful to gain subjective information from participants concerning whether they perceive a trial as memory (i.e., old) or generalisation (i.e., new). Getting such information may allow us to better understand the dorsal striatal and vmPFC activation during memory by ensuring only memory and not generalisation instances are assessed. Using the same analysis procedure used already (i.e., 2x2 GLM), but relying on the subjective judgements of trial type, the relationship between the vmPFC and dorsal striatum could be examined more closely. Specifically, it would be predicted that in trials where participants perceived an item as old, there should be increased BOLD response in the vmPFC and dorsal striatum compared to trials perceived as new.

Presently no differences based on the effect of clustering were observed; this might be due to both conditions being processed as possessing a pattern. However, it is also possible that the effects of clustering may only be observable at the point of encoding as participants learn the word-location associations. For example, studies have shown a shift from retrieval-based to encoding-based representations as more information is accumulated (Bowman et al., 2020). This is also proposed in theories such as SLIMMs (van Kesteren et al., 2012). Given that the clustered condition would be the one to develop a schema due to the overlapping patterns across events (Ghosh & Gilboa, 2014), there may be a shift in the regions responsible for encoding this material. Therefore, if possible, it may be helpful to have participants complete both encoding and retrieval during scanning as a means of assessing changes in regions of activation associated with the task.

4.5.3. Analysis Considerations

Along with analysing the data using the approaches applied here, it may also be useful to assess whether there is evidence of functional coupling between the vmPFC and dorsal striatum during the task. Specifically, functional connectivity analyses could be used to infer whether the two regions are working in tandem based on their correlated BOLD signal over time. This may provide more weight to my speculation that the two regions are working together to prevent interference and maintain only relevant information in working memory.

Future experiments may wish to analyse hippocampal activity based on the anterior and posterior sections instead of the hippocampus as a whole, as was done here. Recent work has found that more anterior portions of the hippocampus are involved in generalisation, whilst the posterior is involved in memory (Bowman & Zeithamova, 2018, 2020; Collin et al., 2015; Frank et al., 2019; Schlichting et al., 2015). As such, in future analyses, it may be worthwhile to divide the long-axis of the hippocampus into the anterior and posterior portions to examine possible functional dissociations. Though subtle, examining the heatmap for the hippocampal ROI (see Figure 4.6), suggests possible evidence of differential recruitment of the region based on trial type, with greater activation for memory trials towards the posterior and reduced activation in the anterior portion. This may explain why no hippocampal effect was found as the small cluster of activation in the posterior region was masked when the average activity across the entire ROI was taken. Therefore, future work should ensure that the anterior and posterior division is used when analysing the hippocampus, particularly when examining memory and generalisation in the same paradigm.

Finally, concerning the lack of clustering effect, it may be helpful to consider methods such as Naïve Bayesian or Support Vector Machine classifiers; this has been done to classify cognitive states in other work (e.g., Lee et al., 2009; Ramasangu & Sinha, 2014). Specifically, providing a subset of the fMRI images for each participant and assessing whether the classifier could differentiate between clustered and non-clustered images. Using these methods may allow us to then identify the neural correlates associated with clustering. If it is the case that the two cannot be differentiated, this may support the proposal that both are perceived as possessing a pattern. If they can be differentiated, it will provide insight into how the two are dissociable at a neural level. One belief, based on previous work, is that the vmPFC may be an important area for distinguishing between the two conditions; this relates to its involvement in developing schematic representations of overlapping events (Bowman & Zeithamova, 2018; Ghosh & Gilboa, 2014; van Kesteren et al., 2012). However, it should be noted that some work indicates that the hippocampus may be capable of creating generalised representations of overlapping events (Schlichting et al., 2015).

4.5.4. Open Questions

Given the points discussed within the chapter, the following questions were developed for the analysis of an independent dataset:

1. Are there changes in the regions active during encoding, shifting from hippocampal to more vmPFC throughout learning?
2. Does the posterior hippocampus show greater activation during memory trials, whilst the anterior shows greater activation during generalisation trials?
3. Is there greater activation in the dorsal striatum and vmPFC during memory trials compared to generalisation?
4. Do the dorsal striatum and vmPFC show evidence of functional connectivity?
5. Can the effect of clustering be differentiated using image classifiers (e.g., Naïve Bayesian or Support Vector Machines)?

4.6. Conclusions

The present chapter examined the neural correlates of memory-based generalisation. First an exploration of how the inclusion of an SCT could impact behaviour in the precision paradigm was undertaken, noting no change due to its inclusion. Subsequently, a preliminary analysis of fMRI data occurred, finding the vmPFC and dorsal striatum to be more active during memory than generalisation. However, no differences were found within the hippocampus. From these results, possible changes to the study's design, including implementing subjective judgments and scanning at both Study (encoding) and Test (retrieval) have been suggested. Further, targeted analyses were proposed. For instance, functional connectivity analysis to assess the relationship between vmPFC and dorsal striatum and image classifiers to distinguish the regions involved in processing clustered and non-clustered information. Overall, the preliminary insight provided by the present chapter may support future work using the precision paradigm to understand memory-based generalisation.

Chapter 5:
General Discussion

Part of the General Discussion were also presented in the preprint: Cockcroft, J. P., Berens, S., Gaskell, M., & Horner, A. J. (2021, August 24). *Schematic information influences memory and generalisation behaviour for schema-relevant and -irrelevant information*. https://doi.org/10.31234/osf.io/nzurq. The content was moved to the present Chapter to increase continuity across chapters.

This thesis had three aims: (1) to explore memory-based generalisation using the precision paradigm, (2) to identify whether schematic information could influence information both relevant and irrelevant to itself, and (3) understand the neural processes of memory-based generalisation. In order to do this, the precision paradigm was used. Here, participants were required to retrieve word-location associations around a circle. A pattern was present in one set of word-locations (i.e., the clustered condition), while the other had no underlying pattern (the non-clustered condition). The presence of a pattern in the clustered condition could lead to a schematic representation being developed (Ghosh & Gilboa, 2014; van Kesteren et al., 2012). As such, this paradigm offered an opportunity to examine how schema presence biased memory and generalisation behaviour. Across three chapters, an exploration of the behavioural, computational, and neural processes involved in memory-based generalisation was undertaken.

5.1. Chapter Summaries

5.1.1. Chapter 2

In Chapter 2, an investigation into how the presence of a pattern (or schema) in one condition influenced behaviour in both the clustered and non-clustered conditions was undertaken. In this chapter, four experiments are presented. The first looked at memory and generalisation behaviour immediately after encoding. The second assessed the same measures but with a 24-hour delay between study and test. The inclusion of a delay period was to allow for potential consolidation to take place and magnify the effects observed in Experiment 1. Subsequently, these two experiments were replicated through two online studies to assess the robustness of the effects found.

One of the most interesting findings from this chapter was how the presence of a pattern affected generalisation behaviour for both relevant (clustered) and irrelevant (non-clustered) information. Specifically, participants showed an "avoidance" behaviour, whereby they were less likely to place non-clustered novel items within the same area of the circle as clustered items, despite those items being equally likely anywhere around the circle. Interestingly, this behaviour did not change as a function of time, showing the behaviour both at immediate test and following a 24-hour delay. Despite the non-clustered items being schema-irrelevant, participants behaviour treated this condition as though it did itself have a pattern present – specifically on the opposite side of the circle. Alongside the avoidance effect, it was observed that participants could extract and use the pattern within the clustered condition to make inferences regarding locations for novel items. Finding immediate generalisation behaviour corroborates other findings within the literature (e.g., Djonlagic et al., 2009; Sweegers & Talamini, 2014).

Examination of the memory measures found that clustered items were more accessible than non-clustered, but reporting accuracy (precision) was greater in the latter than the former condition. There was an effect of delay for both memory measures, whereby fewer items were remembered (accessibility), and greater error was found (i.e., less precision), after a 24-hour delay period; this is not surprising as memory declines rapidly post-learning (Ebbinghaus, 1885; as cited in Postman, 1968). However, finding that generalisation remained stable despite a loss of memory does give weight to the idea of a schematic representation being formed (Ghosh & Gilboa, 2014). Further, finding that memory accessibility benefited from the schematic representation but at the expense of precision corroborates the effects described throughout the literature (Arpit et al., 2017; Berens et al., 2020; Rosenbaum et al., 2009; Sekeres et al., 2016).

5.1.2. Chapter 3

The subsequent chapter (Chapter 3) explored mechanisms that may explain the avoidance behaviour observed in Chapter 2. To do this, computational modelling techniques were used to simulate the data presented in Chapter 2. Each model applied a different mechanism to generate a set of responses for the precision task. From there, the model output was evaluated to understand whether the patterns presented mapped onto the pattern of behaviour observed in Chapter 2. For this chapter, three models were created. The first family of models aimed to assess whether basic forms of encoding or retrieval-based models could produce the avoidance behaviour. Subsequently, the second family of models modulated the retrieval probability of items based on their proximity to the cluster centre. Finally, the third model provided an alternative account for the set of behavioural effects observed in Chapter 2 by applying an interference mechanism at both encoding (proximity-based interference) and retrieval (semantic-based interference).

Unsurprisingly, the first family of models that used basic forms of encoding or retrieval-based generalisation could not reproduce the avoidance behaviour. Therefore, there was a need to consider other factors that may influence how memories are relied upon at the point of generalising. To examine one potential influence, the second family of models was developed. This family of models assessed how modulating retrieval probability based on proximity to the cluster may influence generalisation behaviour. Specifically, the model modulated retrieval probability based on an item's proximity to the cluster centre. For clustered items, higher probabilities were given to those items closer to the cluster centre. In contrast, non-clustered items were given higher probabilities the further from the cluster centre they were. Using this method did result in the avoidance behaviour being present during generalisation. However, it also meant avoidance was present during non-clustered memory trials.

Based on the output of the second set of models, an analysis of the non-clustered memory data from Chapter 2 and Berens et al. (2020) occurred. Across both analyses, there was evidence of an avoidance in non-clustered memory. This finding supported the second set of models output by showing that an avoidance behaviour should be present for memory and generalisation. Albeit, the effect in generalisation is larger than in memory (this is discussed in more detail below). Along with supporting the second model, these results suggested that a pattern in one condition affects behaviour towards items that are irrelevant to the pattern itself regardless of whether the items belonged to memory or generalisation trials.

The final model attempted to provide a more mechanistic explanation for the avoidance effect beyond schema presence. Specifically, the model explored whether an interference mechanism, using both proximity-based interference (i.e., modulating retrieval probability based on proximity of other word-locations) and semantic-based interference (i.e., spreading activation to other semantically related items when retrieving the target item) could produce this behaviour in both memory and generalisation. Using this mechanism, the model replicated the broad patterns of behaviour from Chapter 2, thus providing an alternative interpretation for the findings beyond schema.

5.1.3. Chapter 4

Chapter 4 provided a preliminary analysis of an fMRI dataset using the precision paradigm to explore memory-based generalisation. The aim of the chapter was twofold. First, it aimed to assess whether the inclusion of a semantic categorisation task (SCT) prior to taking part in the precision task influenced the behaviour of participants. From there, it aimed to examine the neural correlates of memory-based generalisation using a pilot sample to identify possible design changes and open questions.

The SCT was implemented to reduce item novelty effects that could lead to hippocampal responding during fMRI (see Kumaran & Maguire, 2007, for a discussion). However, in Chapter 2, participants did not often refer to the superordinate categories used within the task (i.e., human-made and natural). Instead, participants often focused on the basic categories (e.g., plants, stationary). Therefore, the inclusion of the SCT raised a concern that it could alter the behaviour of participants during the precision task due to shifting their perception from basic categories to superordinate. However, no change from the mean values obtained in Chapter 2 was found when implementing the SCT. Therefore, it was used in the fMRI experiment.

For the fMRI task, an investigation into the neural correlates of memory-based generalisation took place. Participants completed the SCT and study phase of the precision task outside of the scanner. In the scanner, participants were required to either remember the location associated with a word or make a best guess. It was found that the vmPFC showed greater activation during memory trials compared to generalisation. The dorsal striatum was also marginally significant, with greater activation during memory than generalisation trials. A consideration of possible design and analysis changes in light of the findings were discussed. For instance, scanning during both encoding and retrieval to assess changes in neural responding for the clustering manipulation, analysing the long axis of the hippocampus based on the anterior and posterior regions to examine divisions of labour for memory (posterior) and generalisation (anterior), or the inclusion of subjective judgements to allow for distinguishing between trials relying on memory and those relying on generalisation.

5.1.4. Chapter Outcomes

Across three chapters, an exploration of memory-based generalisation occurred. At the behavioural level, it was evident that the presence of a pattern in one condition affected members within the condition itself and other irrelevant conditions. This effect could be explained via schema influence – the idea that schema bias memory for events based on congruency with expectations. However, the computational chapter challenged this view and proposed an alternative mechanism to explain this behavioural effect: interference. Under this view, the presence of an avoidance behaviour in the non-clustered condition is the result of proximity- and semantic-based interference, without the need to form a schematic representation of events. The final chapter explored the neural correlates of memory-based generalisation, finding preliminary evidence that the vmPFC and dorsal striatum may play important roles in this behaviour. However, firm conclusions cannot be made given the limited sample size and possible design and analysis considerations.

5.2. Implications and Future Directions

5.2.1. The Precision Paradigm

Chapter 1 discussed why the precision paradigm was used in the present work. Briefly, the paradigm provided the opportunity to get a more sensitive measure of memory-based generalisation as it allowed for the examination of patterns extracted and used by participants. In contrast, most studies used binary tasks to assess memory-based generalisation, such as the face-location paradigm (Sweegers & Talamini, 2014) or associative inference (Carpenter & Schacter, 2017; Kumaran, 2013; Preston et al., 2004; Zeithamova et al., 2012). Though these tasks offer the opportunity to assess whether generalisation is possible, they are limited in their sensitivity to understand the patterns extracted and used by participants, instead focusing on “correct” or “incorrect” responses. Using the precision paradigm in the present work identified ways in which a pattern (or schema) in one condition affected both pattern-relevant and -irrelevant material. This novel finding provides unique insights into how schema may influence behaviour (see the Schema Irrelevance sub-section below for more discussion on this point).

Additionally, the paradigm allowed for assessment of not only memory accessibility (i.e., proportion remembered) which many studies use (e.g., Preston et al., 2004; Sweegers & Talamini, 2014), but also memory precision (i.e., angular error from the target to selected location). Therefore, it was possible to assess how memory and generalisation were affected by schematic representations at a finer level of detail than is typically offered in other paradigms. Consequently, future work may benefit from the continued use of this paradigm to explore both memory and generalisation.

5.2.2. Avoidance Behaviour

Throughout the thesis, evidence has been presented demonstrating the presence of an avoidance behaviour for the non-clustered (schema-irrelevant) condition. This has been shown in both the data collected as part of the thesis and a secondary dataset that used a different set of words and did not assess generalisation (Berens et al., 2020); this effect is therefore robust. However, an open question is what this may mean in the broader context of the literature on schema-based effects.

5.2.2.1. Schema Irrelevance: Control Conditions and Theories

The avoidance results highlight that schematic information can affect memory and generalisation behaviour for schema-relevant and -irrelevant information. Experimentally, these results have implications for studies that use schema-irrelevant information as a control condition (e.g., Frank et al., 2018; Greve et al., 2019), where behaviour is assumed not to be affected by the presence of a schema. These studies show that memory performance is enhanced for schema-congruent and -incongruent information relative to schema-irrelevant information. The present results suggest that the presence of schematic information can bias memory for these irrelevant items. Therefore, schema may not boost memory for relevant information but suppress the retrieval of irrelevant information.

Changes in performance for schema-irrelevant information may have been previously missed due to a lack of appropriate control comparison. For example, in Greve et al. (2019), retrieval of schema-irrelevant items may have been reduced by the presence of schematic information, resulting in what appears to be a schema benefit. Instead, the results may be caused by the presence of a schema biasing (i.e., hindering) the retrieval of schema-irrelevant information. Due to comparing behaviour in the non-clustered condition to that expected of a uniform distribution (representing the distribution of locations expected if no biases were present), the present analysis demonstrated biases for schema-irrelevant information that may have been missed in previous studies. Therefore, future studies should be aware that a schema-irrelevant control condition may not be an appropriate baseline given the present results.

Theoretically, the results provide insight into schema processing. They suggest that schematic information affects memory and generalisation behaviour immediately after encoding for schema-relevant and -irrelevant information in a manner that is not predicted by existing schema theories. For example, in the SLIMMs model (van Kesteren et al., 2012), the predominant focus is on schema-congruence and incongruence and how this may affect neural processing. However, as shown in the present work, schemas also affect processing for schema-irrelevant information. Given that this effect was found across the present work and a secondary analysis, theories must incorporate information on how schemas may bias behaviour for events irrelevant to themselves, not simply support enhanced memory for congruent and incongruent events.

5.2.2.2. Schema vs Interference

Throughout the thesis, it is proposed that schema-driven effects may explain the patterns of behaviour presented. Schema have been shown to bias the recall of events, increasing the number of false alarms and leading to errors in reporting events (Bartlett, 1932; Berens et al., 2020; Brewer & Treyens, 1981; Lew & Howe, 2017). Therefore, in the present thesis, finding that the non-clustered condition showed a bias away from the area of the circle where clustered items were located (i.e., those with a schematic representation) provided evidence consistent with a schema bias for irrelevant information.

An alternative interpretation may be that two schemas were formed – one acting for the clustered and the other the non-clustered condition. Indirect support for this notion may come from the lack of differential activation for the clustering effect in the fMRI pilot (Chapter 4); this suggests that both conditions were treated identically at a neural level. Of course, this is a reverse inference, and firm conclusions cannot be drawn. However, it would seem computationally expensive for two schemas to be formed that act on the same behaviour. Instead, one schema may be formed that identifies when an item does not belong to the cluster, it should be placed elsewhere; this would reduce the computational expense. Although, given the debrief information of participants from Chapter 2, it was clear anecdotally that participants were not explicitly using this strategy. The need for explicit awareness is discussed in more detail below.

Though focusing on schematic effects was central to the thesis, consideration for alternative mechanisms for the avoidance effect were made. For instance, in Chapters 2 and 3, it was highlighted that mechanisms such as mutual exclusivity, base rate neglect or interference might provide alternative interpretations for this effect. Subsequently, in Chapter 3, the interference model developed provided a parsimonious explanation for the avoidance behaviour. Specifically, the model showed that proximity- and semantic-based interference were sufficient to produce an avoidance behaviour in both memory and generalisation. Overall, the model provided a good fit for the data reported in Chapter 2 and demonstrated no need for a schematic representation to produce the observed effects.

Along with providing a parsimonious solution for the data reported in Chapter 2, the interference model may explain some of the behavioural effects observed in other experiments. As mentioned previously, studies have speculated that they have studied schema-based processing using the precision paradigm (e.g., Berens et al., 2020; Tompary et al., 2020). However, the present model proposes that these findings may be better explained via interference. In the case of Tompary et al. (2020), they investigated how schema use changed throughout a 1-week period. In their study, two clusters were used, which appeared 180° away from one another. They found that schema use increased over time, as evidenced by an increased tendency to report items closer to the mean of their cluster. However, the accuracy of this schematic representation declined as memory for the individual items themselves did. More specifically, one-week post-learning there was greater divergence between the patterns presented at study and the ones produced by participants compared to when testing occurred 24-hours post-learning. Therefore, despite an increased reliance on the central tendency across events (i.e., reporting items closer to the mean), the schema itself became increasingly less like the originally learned pattern. As schemas are proposed to be stable representations that function independently of the individual experiences themselves (see Ghosh & Gilboa, 2014 for review), finding that schema use declined at the same rate as memory for the individual items may suggest the use of alternative mechanisms. Instead, the behaviour observed in Tompary et al. (2020) may be driven by retrieval-based mechanisms, similar to those used in the interference model described in Chapter 3. Specifically, location judgements may have been based on a weighting of retrieval probability (determined by location proximity) and spreading activation of related items (determined by semantic proximity). Using these weightings, location judgements may appear to show a bias towards the cluster, with participants seemingly showing more “schema-like” behaviour.

Similarly, the above conclusion about interference rather than schema could be applied to the Brady et al. (2018) study. In Brady et al. (2018) participants learned to associate objects with a particular colour. For instance, lamps with the colour green. The colours associated with an object category were drawn from a von Mises distribution, with the primary colour (e.g., green) being the mean of the distribution and therefore being most likely to be appear with an object category. Other nearby colours (e.g., yellow or blue) would also be associated with this category, but to a lesser degree. It was found that participants would increasingly report the colour of an object as the mean of the category (e.g., lamps being green) despite also appearing in other colours. This finding was interpreted to evidence schematic behaviour, whereby there is an increased tendency to rely on the mean. However, interference resulting from proximity could lead to a similar effect. Specifically, there will be a lower probability of retrieving an objects colour when there is greater overlap among items. When an object is then presented, these lower probabilities would be weighed against the psychophysical similarity of the items (similar to the semantic weighting in the interference model of Chapter 3). The lower probability of retrieval along with the psychophysical similarity among items of a similar category will therefore lead to a bias in reporting items as their mean colour. Therefore, what appears to be a schematic bias may be the result of another mechanism entirely, in this case interference. Schurgin et al. (2020) considered this notion of alternative mechanisms when proposing the TCC model (discussed in Chapter 3). Here, psychophysical similarity and signal detection were used to demonstrate how experiences may be misremembered.

Therefore, alternative mechanisms beyond schema may explain some of the findings reported in the present work and the broader literature. What appeared to be a schema-based process may have resulted from an alternative mechanism (e.g., interference). As items were encoded, their probability of retrieval was reduced if an item neighboured previously encoded items. Subsequently, when retrieving the location for the target item, spreading activation would occur, leading to the retrieval of associated items, which bias the reporting of events. Hence, what appears to be a schema may be more simply explained by interference. A similar challenge of interpretation has been shown in the encoding versus retrieval-based literature, whereby retrieval accounts can explain many of the encoding-based effects (see Nosofsky & Zaki, 2002 and Smith & Minda, 2000).

5.2.2.3. Alternative Mechanisms

Interference was the one mechanism used in Chapter 3 to provide an alternative way of conceptualising the patterns of behaviour presented in Chapter 2. However, other proposals were made in Chapters 2 and 3: mutual exclusivity and base rate neglect. Both of these mechanisms are considered in more detail below.

5.2.2.3.1. Mutual Exclusivity

The mutual exclusivity bias refers to the tendency to assign only one label to an object (e.g., viewing a cup and glass as mutually exclusive labels) and is predominately studied in the context of early language learning (Clark, 1988; Golinkoff et al., 1992). However, there has been a recent interest in adult learning and the mutual exclusivity bias (Lake et al., 2019). Concerning the results reported within the present thesis, one area of the circle will be perceived as devoted to the clustered category, whilst the other would represent the non-clustered. Therefore, there will be a tendency to report locations for non-clustered items on the opposite side of the circle. This would explain the presence of an avoidance effect in both memory and generalisation.

The simplest version of a mutual exclusivity bias might be an explicit process at retrieval where, when a new non-clustered word is presented, participants actively retrieve a schema related to the clustered condition and use an “if not in the clustered category, place on the opposite side of the circle” strategy. Although possible, analysis of the post-retrieval debrief suggests very few participants were using explicit strategies such as these. Further, the mutual exclusivity bias would require participants to be categorising items at the superordinate level (i.e., human-made and natural). The problem with this assumption is that the debrief suggested few participants in Chapter 2 (4 of 261, 1.53%) spontaneously referred to items belonging to these semantic groupings. Instead, participants were more likely to categorise words into basic categories (e.g., “household objects, animals and fruit”, “fruit and vegetables, household items, mammals”, “fruit, technology… cars… exotic animals, weather”, “planets, animals, food”). Of course, it is not possible to rule out that the mutual exclusivity bias worked implicitly (Merriman & Stevenson, 1997), meaning the distinction between human-made and natural items occurred without conscious awareness.

If a mutual exclusivity bias was causing the effect, one critical question is what is driving this bias? One explanation would be that the non-clustered condition is unlike most groupings found in the real world. Returning to the bathroom example described in Chapters 1 and 2, if you have a “bathroom” schema, it is probably less likely that you will find non-bathroom related items in this location relative to elsewhere in the house. In short, the “bathroom” schema does not tell you where the microwave will be, but it likely provides information about where it is unlikely to be. Although real-world examples of more uniformly distributed items may exist (e.g., house plants throughout a home), they may be rare, and as such, we may have little experience with them. Participants may apply this real-world sampling experience to the present experiments, presuming non-clustered words are less likely to be located in the clustered area of the circle.

5.2.2.3.2. Base-Rate Neglect

In contrast, base rate neglect would propose that participants will be biased by the relative rather than absolute probabilities of events when generalising (Hawkins et al., 2015; Welsh & Navarro, 2012; Wolfe, 2007). Specifically, though non-clustered items had an absolute probability of appearing anywhere around the circle, their relative probability of appearing in the clustered area of the circle compared to clustered items was much smaller. As such, participants may rely on this perceived difference in relative probability to infer locations during generalisation. Since memory items can either rely on a remembered instance (i.e., the target location for a given item) or generalisation when an item is not remembered, the avoidance behaviour will be less apparent during memory compared to generalisation as generalisation mechanisms will only be relied upon if an item is not remembered. In contrast, generalisation trials cannot rely on memory for the location associated with the novel item. These trials rely only on the relative probability of events leading to greater avoidance behaviour. Therefore, unlike the mutual exclusivity bias, the base rate neglect proposal may be more able to explain the larger presence of avoidance during generalisation.

One way of potentially testing this theory could be the use of subjective judgements. Specifically, for each trial, participants could identify whether an item was perceived as “old” or “new”. Doing this will allow for a clear dichotomy between episodes relying on a memory for an individual episode or when they generalised to make a location judgement. Under the base-rate neglect proposal, when items are perceived as "old" (or "remembered"), there should be uniformity present in the locations for non-clustered items. In contrast, when the item is perceived as non-remembered (or novel), the locations will be selected based on generalisation mechanisms, thus leading to avoidance for the clustered area.

Another advantage of the base-rate neglect proposal is that it can explain the avoidance behaviour without considering the semantic groupings at the superordinate level. Whilst it would be necessary for the mutual exclusivity bias to consider the groupings at the superordinate level, the base rate neglect proposal could function so long as participants were sensitive to the semantic distances between words. Similar to the interference-based model proposed in Chapter 3, the base-rate neglect proposal does not require a clear dichotomy between clustered and non-clustered items. Instead, the model could use the semantic distances between items to dissociate between semantic groupings. As such, the base rate neglect proposal could explain the avoidance effect without categorisation at the superordinate level of human-made and natural items. However, given research has suggested that base-rate neglect is driven by explicit processes (Lovett & Schunn, 1999; c.f. Bohil & Wismer, 2015; Wismer & Bohil, 2017), it is possible that the avoidance behaviour observed throughout the thesis (if base-rate neglect is the correct explanation) would be sensitive to whether participants are learning word-location associations under conditions that preclude explicit awareness. The need for explicit awareness is discussed in more detail below.

Regardless of the perceived mechanism, an open question is what would occur if encoding were to take place over several days rather than in one session. Would an avoidance behaviour still be observed? The reduced memory load caused by encoding over several days may lead to a loss of avoidance behaviour for the interference model. In contrast, schema-driven biases may predict that the avoidance behaviour would remain given non-clustered items would still appear within the clustered area "unexpectedly" and therefore have a reduced probability of retrieval. Additionally, if learning occurred over several days, a schematic representation may be more likely to develop. In their rodent study, Richards et al. (2014) had mice learn platform locations that followed an underlying pattern. This learning took place over several days and several repetitions of the same platform. It was found that mice 30 days post-training would swim more around the average location than elsewhere, suggesting the development of a schematic representation. Suppose a similar training approach of spaced learning had taken place for the behavioural work reported in Chapter 2. In that case, the avoidance behaviour observed may (assuming a schema-based mechanism) or may not have been present (assuming an interference mechanism). Future work may benefit in assessing whether the avoidance effect would be observed when encoding occurs over time instead of the mass encoding presented in Chapter 2.

5.2.2.3.3. Summary

Overall, it is clear that several different interpretations for the avoidance behaviour could be given – from schema, interference, mutual exclusivity and base-rate neglect. Across chapters, it appears that schema-driven mechanisms may not be the sole explanation for the avoidance behaviour found. Chapter 3 identified that interference mechanisms might provide a more parsimonious solution for the presence of an avoidance behaviour. Though a reverse inference, one interpretation for the results of Chapter 4 was that two schemas may have been formed for the clustered and non-clustered conditions, respectively. Additionally, as discussed above, mutual exclusivity and base-rate neglect also give practical interpretations for the avoidance effect. Future work may benefit from including subjective judgments to disentangle the base-rate neglect proposal that avoidance would only be observed when participants do not remember the location associated with an item or perceive it to be novel.

5.2.2.4. Explicit or implicit processes

An open question is whether the behaviour presented during the thesis resulted from explicit awareness of the pattern. Though data was collected via the Introspection Questionnaire as to whether participants perceived a pattern in the way items were presented, no formal analysis was undertaken as no specific hypotheses were drafted at the time of writing. However, a question that has arisen throughout the work is whether explicit or implicit mechanisms drive these effects. Specifically, does the avoidance behaviour require an explicit awareness of the clustered pattern? In Chapter 2, 152 (58.24%) of the 261 participants stated that they perceived a pattern in the word-locations presented. Given not all participants show the avoidance effect, if an analysis was undertaken to assess whether an avoidance behaviour was present only when a pattern was perceived, this might suggest that explicit rather than implicit mechanisms drive this effect. From this, further weight may be given to the base-rate neglect proposal, whereby an explicit awareness of the pattern would be required.

Along with this, it may be worthwhile for future work to consider how to increase the perception of a pattern in the present paradigm. Presently, the paradigm used within the thesis requires a minimum of 70 memory and 15 generalisation items to generate reliable estimates of memory and generalisation performance. As such, at a minimum, 85 words per condition are required for the present task. The consequence of this is using superordinate categories of words (i.e., human-made and natural). However, using the superordinate categories can be difficult, with many studies showing participants preference for using basic categories (e.g., fruit and furniture; Rosch et al., 1976). The differentiation between basic-level categories is much easier to accomplish as members within a category are very similar compared to members of other categories (Murphy & Brownell, 1985). As a result, it is easier to distinguish between category members when the category is more cohesive. Though anecdotal, of the 152 participants that perceived a pattern in the word-locations for Chapter 2, 82 (53.95%) perceived a pattern when natural items were clustered compared to 70 (46.05%) when human-made items were clustered. This suggests that patterns may have been easier to extract for natural items, which are more cohesive than human-made items.

Had a more cohesive set of items been used (e.g., fruit and furniture), is it possible that the avoidance effect would have been exacerbated as a clear dichotomy between groups could be made? Of course, this would likely require the implementation of different analysis procedures. Presently, mixture model estimation is used to extract the p and k memory measures, where a minimum of 70 items per condition is required to get a reliable estimate of memory performance. Hence why the present work used superordinate categories of human-made and natural. However, other studies using the precision paradigm have simply used the angular error of responses to estimate memory accuracy (e.g., Tompary et al., 2020). Doing this would prevent the fitting issues associated with the mixture model whilst allowing for the analysis of memory and generalisation behaviour. Using this analysis procedure, where fewer exemplar would be required, means that a more cohesive set of items (e.g., fruit and furniture) could be used. Moving from the superordinate (i.e., human-made and natural) to basic categories may exacerbate the avoidance behaviour. Although, it is worth bearing in mind that this would result in no longer being able to examine differences in memory accessibility and precision simultaneously – one of the advantages offered via the present paradigm. Therefore, alternative ways of making the categories more prominent may be required should one want to examine both accessibility and precision simultaneously. For example, using different fonts or colours for each category during learning.

5.2.3. Generalisation and Consolidation Models

In Chapter 2, generalisation of clustered items to the underlying pattern (i.e., the von Mises distribution) was observed immediately and remained stable after a 24-hour delay. This is in line with some research showing that generalisation ability remains relatively stable over more extended periods (Ellenbogen et al., 2007; Mirković & Gaskell, 2016; Sweegers & Talamini, 2014; Tamminen et al., 2012). Though, in some of these investigations, there was a demonstration of initial improvement in generalisation ability. For instance, Sweegers and Talamini (2014) found that generalisation from immediate test to 4-hours post-study significantly increased. However, after a 1-month delay, generalisation ability remained the same. Though the initial finding of change was not demonstrated in the present work, this may be due to the time-points at which testing occurred and the sensitivity of the precision paradigm.

These results are contrary to other findings that show generalisation capacity declines over time (Tompary et al., 2020). The discrepancy between the present thesis results and those of Tompary et al. (2020) is likely also related to the time points examined. In Tompary et al. (2020), generalisation was assessed at 24-hours and 1-week post-encoding. In contrast, the present thesis assessed generalisation immediately and 24-hours later. Given the extended delay period in the Tompary et al. (2020), the finding of a change in generalisation may have occurred. If a similar level of delay was used in the present work, a similar decline in generalisation capacity might have been observed. Future work may benefit from looking at longer delays; this may allow for better examination of whether a schematic representation was developed given they should remain stable over time (Ghosh & Gilboa, 2014; van Kesteren et al., 2012). Though, if no decline in generalisation capacity is found, this would be contrary to the findings of Tompary et al. (2020), which may further corroborate the earlier arguments that they were not examining schema-based processes.

Finding immediate generalisation also poses a challenge for some systems consolidation models. Specifically, according to the CLS (McClelland et al., 1995) model, generalisation would require the extraction of commonalities across experiences, driven by consolidation (i.e., the slower learning system); this can take hours, days, weeks or years (Rasch & Born, 2013). However, the present thesis demonstrated immediate generalisation without consolidation; this challenges the proposal made by CLS. Here, participants were able to extract and use the commonalities across experiences to generalise to novel instances without needing a delay period; this ability also showed no change following a period of consolidation. Therefore, the present results pose a challenge for the CLS model.

However, more recent extensions of the CLS model, such as REMERGE (Kumaran & McClelland, 2012), have overcome the issue described. Specifically, REMERGE proposes that generalisation may be possible immediately via the hippocampus (i.e., the "fast learning system"). Under this proposal, the hippocampus can use retrieval-based mechanisms to allow for generalisation to novel instances. Therefore, the slower extraction of commonalities across experiences is not required for generalisation to take place. Consequently, finding that generalisation did occur immediately with little change over time is more in line with the REMERGE model, which extended the CLS proposals.

5.2.4. Neural Mechanisms of Memory-Based Generalisation

Chapter 4 aimed at exploring the neural mechanisms of memory-based generalisation. However, due to the COVID-19 pandemic, the study could not continue. As a result, a preliminary analysis was conducted on the existing data to identify future hypotheses, design changes and analysis recommendations for a future study. As described in the chapter itself, the vmPFC and dorsal striatum may play important roles in the present task. It was hypothesised this may be due to functional coupling between the two regions as they attempted to gate working memory (the dorsal striatum) and prevent interference (the vmPFC; O'Reilly & Frank, 2006). Surprisingly, however, no differential activation of the hippocampus was found. This is contrary to a plethora of studies that find the hippocampus to be important to generalisation (Bowman & Zeithamova, 2018; Frank et al., 2019; Schlichting et al., 2015). It was speculated this may be due to a lack of statistical power or due to the way in which the hippocampus had been analysed. Specifically, not splitting the hippocampus into the anterior and posterior portions. Research has shown that posterior regions are associated with memory-based processes, whilst anterior portions are associated with generalisation (Schlichting et al., 2015). Future work may benefit from considering some open questions. For instance, are the vmPFC and dorsal striatum functionally coupled, allowing for memory gating of irrelevant information? Would the inclusion of "old" or "new" judgements allow for examination of this hypothesis? Can the effect of clustering be differentiated within the brain? Does the hippocampus differentially recruit anterior and posterior regions based on task demands?

5.3. Conclusion

This thesis aimed at exploring memory-based generalisation using the precision paradigm to explore how schematic information influenced behaviour. Across three chapters that employed different techniques (i.e., behavioural, computational and neuroimaging) it was observed that schema-based processes might influence behaviour for both schema-relevant and -irrelevant information. This poses questions for studies implementing a schema-irrelevant control condition and many theories that overlook schema-irrelevant information in their conceptualisations. However, there are alternative interpretations that could be posed to the avoidance behaviour found. For instance, interference-based mechanisms may drive the avoidance behaviour without schematic influence, as shown via computational modelling. This alternative interpretation of the findings could also be applied to other studies and raises questions about whether these studies have assessed schema-based processing. Along with interference, mutual exclusivity and base-rate neglect were alternative proposals applied to the behaviour observed in the present thesis. Future modelling or experimental work may be required to disentangle which of these explanations provides the best fit for the data. Nevertheless, the thesis provides clear evidence that the presence of a pattern can affect both memory and generalisation for both relevant and irrelevant information.

Appendices

Table of Contents

Appendix A:

Contrast Matrices for Experimental Effects

Throughout the thesis, dummy-coding was used to construct the GLME's. A consequence of this is that the output from the model, when higher-order terms are present, will refer to simple effects as opposed to the effects of interest. Therefore, to ensure the relevant hypotheses were tested, linear hypothesis tests were conducted on the model coefficients via the `coefTest` function in MATLAB Statistics and Machine Learning Toolbox. This function allows for specific tests to be performed on the model coefficients. Below are the contrast matrices used for each of the analyses where higher-order terms were present; this is split by Chapter. Note, when looking at the avoidance effect, a comparison is made from the average value irrespective of parameters against the uniform density ($2\pi^{-1}$).

Chapter 2

In Chapter 2, two forms of GLME were computed in order to assess the memory and generalisation measures. The first set of models assessed the effects of Clustering (0 = Non-Clustered, 1 = Clustered), Delay (0 = Immediate Test, 1 = Delayed Test) and Setting (0 = In-Lab, 1 = Online) on total information, accessibility, precision and D_{KL} von Mises. Below is a table (Table A1) of contrast matrices used for the comparisons reported.

Table A1.
Contrast matrices for the analysis of total information, accessibility, precision and D_{KL} von Mises.

Hypothesis	Parameter							
	I	**C**	**D**	**S**	**C x D**	**C x S**	**D x S**	**C x D x S**
Clustering: Is there a main effect of clustering?	0	1	0	0	0.5	0.5	0	0.25
Delay: Is there a main effect of delay?	0	0	1	0	0.5	0	0.5	0.25

Setting: Is there a main effect of setting?	0	0	1	0	0	0.5	0.5	0.25
Clustering x Delay: Does the effect of clustering change over time?	0	0	0	0	1	0	0	0.5
Clustering x Setting: Does the effect of clustering change by setting?	0	0	0	0	0	1	0	0.5
Delay x Setting: Does the effect of delay differ across settings?	0	0	0	0	0	0	1	0.5
Clustering x Delay x Setting: Does clustering change as a function of delay and setting?	0	0	0	0	0	0	0	1

Note. *I* = Intercept, *C* = Clustering, *D* = Delay, *S* = Setting.
These values are entered into the `coefTest` function from MATLAB Statistics and Machine Learning Toolbox to derive the *t*, *p* and degrees of freedom for each contrast.

For the GLME that assessed memory precision, a Clustering x Setting interaction was found to be significant. To explore this further, post-hoc analyses were conducted. Table A2 provides the matrices used to explore these effects.

Table A2.
Contrast matrices for the post-hoc analysis of Clustering x Setting.

Comparison	**Parameter**							
	I	**C**	**D**	**S**	**C x D**	**C x S**	**D x S**	**C x D x S**
Non-Clustered In-Lab vs. Clustered In-Lab	0	-1	0	0	-0.5	0	0	0
Non-Clustered In-Lab vs. Non-Clustered Online	0	0	0	-1	0	0	-0.5	0
Non-Clustered In-Lab vs. Clustered Online	0	-1	0	-1	-0.5	-1	-0.5	-0.5
Clustered In-Lab vs. Non-Clustered Online	0	1	0	-1	0.5	0	-0.5	0
Clustered In-Lab vs. Clustered Online	0	0	0	-1	0	-1	-0.5	-0.5
Non-Clustered Online vs. Clustered Online	0	-1	0	0	-0.5	-1	0	-0.5

Note. *I* = Intercept, *C* = Clustering, *D* = Delay, *S* = Setting.

These values are entered into the `coefTest` function from MATLAB Statistics and Machine Learning Toolbox to derive the *t*, *p* and degrees of freedom for each contrast.

The second form of GLME examined the avoidance behaviour in generalisation. For this model, only the effects of Delay and Setting were input. Below is a table (Table A3) of contrast matrices used to assess the effects of interest.

Table A3.
Contrast matrices for the kernel density estimates of non-clustered trials.

Hypothesis	**Parameter**			
	I	**D**	**S**	**D x S**
Avoidance: Is there evidence of avoidance behaviour?	1	0.33	0.33	0.33
Delay: Is there a difference in avoidance over time?	0	1	0	0.5
Setting: Does the avoidance behaviour change across settings?	0	0	1	0.5
Delay x Setting: Does the avoidance behaviour change differentially based on delay and setting?	0	0	0	1

Note. *I* = Intercept, *D* = Delay, *S* = Setting.
These values are entered into the `coefTest` function from MATLAB Statistics and Machine Learning Toolbox to derive the *t*, *p* and degrees of freedom for each contrast.

Chapter 3

In Chapter 3, an exploration of the Trial Type (0 = Memory, 1 = Generalisation), Delay (0 = Immediate Test, 1 = Delayed Test) and Setting (0 = In-Lab, 1 = Online) was undertaken to explore non-clustered avoidance in memory along with differences in avoidance across memory and generalisation trials. This was conducted as an exploratory analysis within this chapter. Table A4, below, provides the contrast matrices computed to assess the effects of interest.

Table A4.
Contrast matrices for the exploratory analysis of avoidance in memory.

Hypothesis	Parameter							
	I	**T**	**D**	**S**	**T x D**	**T x S**	**D x S**	**T x D x S**
Avoidance: Do the densities differ from uniform?	1	0	0.5	0.5	0	0	0.25	0
Delay: Do densities in memory change over time?	0	0	1	0	0	0	0.5	0
Setting: Do densities in memory change based on setting?	0	0	0	1	0	0	0.5	0
Delay x Setting: Do delay and setting interact isolated to memory?	0	0	0	0	0	0	1	0
Trial Type: Do memory and generalisation differ?	0	1	0	0	0.5	0.5	0	0.25

Note. *I* = Intercept, *T* = Trial Type, *D* = Delay, *S* = Setting.
These values are entered into the `coefTest` function from MATLAB Statistics and Machine Learning Toolbox to derive the *t*, *p* and degrees of freedom for each contrast.

Along with the above, a confirmatory analysis was undertaken on the Berens et al. (2020) dataset to explore whether there was also evidence of avoidance that did not change as a function of delay. Here, there were five delay periods used from 0hrs to 24hrs. Table A5 provides the contrast matrices used to assess the effects of interest.

Table A5.
Contrast matrices for the confirmatory analysis of avoidance in memory.

Hypothesis	Parameter				
	I	**D: 3hrs**	**D: 6hrs**	**D: 12hrs**	**D: 24hrs**
Avoidance: Do the densities differ from uniform?	1	0.20	0.20	0.20	0.20
Delay: Do densities change over time?	0	1	0	0	0
	0	0	1	0	0
	0	0	0	1	0
	0	0	0	0	1

Note. *I* = Intercept. All other values represent the individual delay periods. These values are entered into the `coefTest` function from MATLAB Statistics and Machine Learning Toolbox to derive the *t*, *p* and degrees of freedom for each contrast.

Chapter 4

For the behavioural pilot in this chapter, an analysis was undertaken to estimate whether the effect of clustering remained the same between Chapters 2 and 4 for each of the effects of interest. Two types of GLME were computed. The first set of GLME's examined the effects of total information, accessibility, precision and D_{KL} von Mises. For these models, the following parameters were entered: Clustering (0 = Non-Clustered, 1 = Clustered), Delay (0 = Immediate Test, with 3-hours and 24-hours explicitly included in the model) and Setting (0 = In-Lab, 1 = Online). Clustering was made to interact with Delay and Setting independently due to a lack of 3-hour Online effect. Therefore, Delay and Setting did not interact in these models. Table A6, below, provides the contrast matrices used to examine the effects of interest for these four models.

Table A6.
Contrast matrices for the analysis of: total information, accessibility, precision and D_{KL} von Mises.

Hypothesis	**Parameter**							
	I	**C**	**D: 3hrs**	**D: 24hrs**	**S**	**C x D: 3hrs**	**C x D: 24hrs**	**C x S**
Clustering (Pilot): Does the pilot study show an effect of clustering?	0	-1	0	0	0	-1	0	0
Delay: Is there a difference in the estimate derived for Chapter 2 and the pilot?	0	0	1	-0.5	-0.5	0.5	-0.25	-0.25
Clustering x Delay: Does the effect of clustering differ between Chapter 2 and the pilot?	0	0	0	0	0	1	-0.5	-0.5

Note. *I* = Intercept, *C* = Clustering, *D: 3-hrs* = Delay at 3-hrs, *D: 24-hrs* = Delay at 24-hrs, *S* = Setting.

These values are entered into the `coefTest` function from MATLAB Statistics and Machine Learning Toolbox to derive the *t*, *p* and degrees of freedom for each contrast.

A further two GLME's were computed, but did not include the Clustering effect as a parameter. Instead, the analyses only included Delay and Setting as the effects of interest were isolated to the non-clustered condition. Specifically, an exploration of avoidance behaviour was conducted. Table A7 provides the contrast matrices used to explore the effects of interest.

Table A7.
Contrast matrices for the kernel density estimates of non-clustered trials.

Hypothesis	**Parameter**			
	I	**D: 3hrs**	**D: 24hrs**	**S**
Avoidance (Pilot): Is there evidence of avoidance behaviour in the pilot data?	1	1	0	0
Delay: Is there a difference in the estimate derived for Chapter 2 and the pilot?	0	1	-0.5	-0.5

Note. *I* = Intercept, *D: 3-hrs* = Delay at 3-hrs, *D: 24-hrs* = Delay at 24-hrs, *S* = Setting.
These values are entered into the `coefTest` function from MATLAB Statistics and Machine Learning Toolbox to derive the *t*, *p* and degrees of freedom for each contrast.

Appendix B:

Retrieval Probability Function – The Equations

The retrieval probability function was developed by Sam Berens (S. C. Berens, personal communication, 20 August 2020) to determine the probability of retrieving a location given its proximity to an arbitrary angle. Below is an in-depth overview of the mathematics behind how this function works.

First, we define a sigmoidal function denoted ($\mathrm{Pr}_{\mathrm{ret}}(\,\theta \mid a, b\,)$) that provided a bounded probability value (i.e., between 0 and 1) describing the likelihood that a location would be retrieved from memory based on its proximity to a given angle (*ϑ*). Equation 1, below, shows this sigmoidal function.

$$\mathrm{Pr}_{\mathrm{ret}}(\,\theta \mid a, b\,) = \frac{1}{1 + \exp(a \cdot g(\theta) + b)} \tag{1}$$

The parameter a was chosen to satisfy the integral (see 2), where p denoted the prior probability of retrieval success across the entire domain of the circle (-π, π).

$$\int_{-\pi}^{\pi} f(\theta) \cdot \mathrm{Pr}_{\mathrm{ret}}(\,\theta \mid a, b\,)\, d\theta = p \tag{2}$$

b was calculated (see 3) to determine the minimum probability value that could be returned by $\mathrm{Pr}_{\mathrm{ret}}(\,\theta \mid a, b\,)$. *Ψ* was the free parameter that determined this minimum probability. Across simulations, this value was set to 0.01.

$$b = \log\left(\frac{1 - \psi}{\psi}\right) \tag{3}$$

Finally, $g(\theta)$ represented the standard von Mises probability density function (see 4), where *μ* denoted the mean location of the cluster, *κ* the concentration and $I_0(\kappa)$ the modified Bessel function of the first kind with order 0 evaluated at the point of *κ*.

$$g(\theta) = \frac{e^{\kappa \cos(\theta - \mu)}}{2\pi I_0(\kappa)} \tag{4}$$

Using these parameters, the sigmoidal function ($\mathrm{Pr}_{\mathrm{ret}}(\theta \mid a, b)$) was able to provide an estimated probability of an item being remembered based on its proximity to an angular location (e.g., -π).

References

Aggleton, J. P., & Brown, M. W. (1999). Episodic memory, amnesia, and the hippocampal–anterior thalamic axis. *Behavioral and Brain Sciences*, *22*(3), 425–444. https://doi.org/10.1017/S0140525X99002034

Aizpurua, A., Garcia-Bajos, E., & Migueles, M. (2009). False memories for a robbery in young and older adults. *Applied Cognitive Psychology*, *23*(2), 174–187. https://doi.org/10.1002/acp.1461

Albert, M. S., Butters, N., & Levin, J. (1979). Temporal gradients in the retrograde amnesia of patients with alcoholic Korsakoff's disease. *Archives of Neurology*, *36*(4), 211–216. https://doi.org/10.1001/archneur.1979.00500400065010

Anderson, M. C., & Neely, J. H. (1996). Interference and Inhibition in Memory Retrieval. In E. L. Bjork & R. A. Bjork (Eds.), *Memory* (pp. 237–313). Academic Press. https://doi.org/10.1016/B978-012102570-0/50010-0

Anderson, R. C. (1984). Some Reflections on the Acquisition of Knowledge. *Educational Researcher*, *13*(9), 5–10. https://doi.org/10.3102/0013189X013009005

Antony, J. W., Ferreira, C. S., Norman, K. A., & Wimber, M. (2017). Retrieval as a Fast Route to Memory Consolidation. *Trends in Cognitive Sciences*, *21*(8), 573–576. https://doi.org/10.1016/j.tics.2017.05.001

Antony, J. W., Stiver, C. A., Graves, K. N., Osborne, J., Turk-Browne, N. B., & Bennion, K. A. (2021). Spatial gist extraction during human memory consolidation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. https://doi.org/10.1037/xlm0000894

Arpit, D., Jastrzębski, S., Ballas, N., Krueger, D., Bengio, E., Kanwal, M. S., Maharaj, T., Fischer, A., Courville, A., Bengio, Y., & Lacoste-Julien, S. (2017). A Closer Look at Memorization in Deep Networks. *Proceedings of the 34th International Conference on Machine Learning*, 233–242. https://proceedings.mlr.press/v70/arpit17a.html

Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *NeuroImage*, *38*(1), 95–113. https://doi.org/10.1016/j.neuroimage.2007.07.007

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

Baddeley, A., & Hitch, G. J. (1977). Recency re-examined. In S. Dornic (Ed.), *Attention and Performance VI* (pp. 647–667). Lawrence Erlbaum Associates.

Barbeau, E. J., Chauvel, P., Moulin, C. J. A., Regis, J., & Liégeois-Chauvel, C. (2017). Hippocampus duality: Memory and novelty detection are subserved by distinct mechanisms. *Hippocampus*, *27*(4), 405–416. https://doi.org/10.1002/hipo.22699

Barnes, J. M., & Underwood, B. J. (1959). 'Fate' of first-list associations in transfer theory. *Journal of Experimental Psychology*, *58*(2), 97–105. https://doi.org/10.1037/h0047507

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

Batterink, L. J., & Paller, K. A. (2017). Online neural monitoring of statistical learning. *Cortex*, *90*, 31–45. https://doi.org/10.1016/j.cortex.2017.02.004

Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, *9*(10), 1–11. https://doi.org/10.1167/9.10.7

Beatty, W. W., Salmon, D. P., Butters, N., Heindel, W. C., & Granholm, E. L. (1988). Retrograde amnesia in patients with Alzheimer's disease or Huntington's disease. *Neurobiology of Aging*, *9*(2), 181–186. https://doi.org/10.1016/S0197-4580(88)80048-4

Berens, P. (2009). CircStat: A MATLAB Toolbox for Circular Statistics. *Journal of Statistical Software*, *31*(10), 1–21. https://doi.org/10.18637/jss.v031.i10

Berens, S. C. (2020, August 20). *Retrieval probability function* [Personal Communication].

Berens, S. C., & Bird, C. M. (2021). *Medial temporal and prefrontal cortices encode structural task representations at different levels of abstraction*. https://doi.org/10.1101/2021.08.11.455977

Berens, S. C., Richards, B. A., & Horner, A. J. (2020). Dissociating memory accessibility and precision in forgetting. *Nature Human Behaviour*, *4*(8), 866–877. https://doi.org/10.1038/s41562-020-0888-8

Best, D. J., & Fisher, N. I. (1979). Efficient Simulation of the von Mises Distribution. *Journal of the Royal Statistical Society*, *28*(2), 152–157. https://doi.org/10.2307/2346732

Bohil, C. J., & Wismer, A. J. (2015). Implicit learning mediates base rate acquisition in perceptual categorization. *Psychonomic Bulletin & Review*, *22*(2), 586–593. https://doi.org/10.3758/s13423-014-0694-2

Bower, G. H., Black, J. B., & Turner, T. J. (1979). Scripts in memory for text. *Cognitive Psychology*, *11*(2), 177–220. https://doi.org/10.1016/0010-0285(79)90009-4

Bowman, C. R., Iwashita, T., & Zeithamova, D. (2020). Tracking prototype and exemplar representations in the brain across learning. *ELife*, *9*. https://doi.org/10.7554/eLife.59360

Bowman, C. R., & Zeithamova, D. (2018). Abstract Memory Representations in the Ventromedial Prefrontal Cortex and Hippocampus Support Concept Generalization. *The Journal of Neuroscience*, *38*(10), 2605–2614. https://doi.org/10.1523/JNEUROSCI.2811-17.2018

Bowman, C. R., & Zeithamova, D. (2020). Training set coherence and set size effects on concept generalization and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(8), 1442–1464. https://doi.org/10.1037/xlm0000824

Brady, T., Schacter, D. L., & Alvarez, G. (2018). *The adaptive nature of false memories is revealed by gist-based distortion of true memories*. PsyArXiv. https://doi.org/10.31234/osf.io/zeg95

Brett, M., Anton, J. L., Valabregue, R., & Poline, J. B. (2002, June 2). *Region of interest analysis using an SPM toolbox* [Conference]. Eighth International Conference on Functional Mapping of the Human Brain, Sendai, Japan.

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

Brown, M. W., & Aggleton, J. P. (2001). Recognition memory: What are the roles of the perirhinal cortex and hippocampus? *Nature Reviews Neuroscience*, *2*(1), 51–61. https://doi.org/10.1038/35049064

Cann, D. R., McRae, K., & Katz, A. N. (2011). False Recall in the Deese–Roediger–McDermott Paradigm: The Roles of Gist and Associative Strength. *Quarterly Journal of Experimental Psychology*, *64*(8), 1515–1542. https://doi.org/10.1080/17470218.2011.560272

Carpenter, A. C., & Schacter, D. L. (2017). Flexible retrieval: When true inferences produce false memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(3), 335–349. https://doi.org/10.1037/xlm0000340

Cipolotti, L., Shallice, T., Chan, D., Fox, N., Scahill, R., Harrison, G., Stevens, J., & Rudge, P. (2001). Long-term retrograde amnesia… the crucial role of the hippocampus. *Neuropsychologia*, *39*(2), 151–172. https://doi.org/10.1016/S0028-3932(00)00103-2

Clark, E. V. (1988). On the logic of contrast. *Journal of Child Language*, *15*(2), 317–335. https://doi.org/10.1017/S0305000900012393

Collin, S. H. P., Milivojevic, B., & Doeller, C. F. (2015). Memory hierarchies map onto the hippocampal long axis in humans. *Nature Neuroscience*, *18*(11), 1562–1564. https://doi.org/10.1038/nn.4138

Cooper, R. A., & Ritchey, M. (2020). Progression from Feature-Specific Brain Activity to Hippocampal Binding during Episodic Encoding. *The Journal of Neuroscience*, *40*(8), 1701–1709. https://doi.org/10.1523/JNEUROSCI.1971-19.2019

Crowder, R. G. (1976). *Principles of learning and memory*. Wiley.

Cycowicz, Y. M., Nessler, D., Horton, C., & Friedman, D. (2008). Retrieving object color: The influence of color congruity and test format. *NeuroReport: For Rapid Communication of Neuroscience Research*, *19*(14), 1387–1390. https://doi.org/10.1097/WNR.0b013e32830c8df1

Dale, A. M. (1999). Optimal experimental design for event-related fMRI. *Human Brain Mapping*, *8*(2–3), 109–114. https://doi.org/10.1002/(SICI)1097-0193(1999)8:2/3<109::AID-HBM7>3.0.CO;2-W

Davis, T., Love, B. C., & Preston, A. R. (2012). Learning the Exception to the Rule: Model-Based fMRI Reveals Specialized Representations for Surprising Category Members. *Cerebral Cortex*, *22*(2), 260–273. https://doi.org/10.1093/cercor/bhr036

Davis, T., & Poldrack, R. A. (2014). Quantifying the Internal Structure of Categories Using a Neural Typicality Measure. *Cerebral Cortex*, *24*(7), 1720–1737. https://doi.org/10.1093/cercor/bht014

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*(1), 17–22. https://doi.org/10.1037/h0046671

Dennis, N. A., Kim, H., & Cabeza, R. (2008). Age-related Differences in Brain Activity during True and False Memory Retrieval. *Journal of Cognitive Neuroscience*, *20*(8), 1390–1402. https://doi.org/10.1162/jocn.2008.20096

DeVito, L. M., Lykken, C., Kanter, B. R., & Eichenbaum, H. (2010). Prefrontal cortex: Role in acquisition of overlapping associations and transitive inference. *Learning & Memory*, *17*(3), 161–167. https://doi.org/10.1101/lm.1685710

Djonlagic, I., Rosenfeld, A., Shohamy, D., Myers, C., Gluck, M., & Stickgold, R. (2009). Sleep enhances category learning. *Learning & Memory*, *16*(12), 751–755. https://doi.org/10.1101/lm.1634509

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

Ellenbogen, J. M., Hu, P. T., Payne, J. D., Titone, D., & Walker, M. P. (2007). Human relational memory requires time and sleep. *Proceedings of the National Academy of Sciences*, *104*(18), 7723–7728. https://doi.org/10.1073/pnas.0700094104

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

Fernández, G., & Morris, R. G. M. (2018). Memory, Novelty and Prior Knowledge. *Trends in Neurosciences*, *41*(10), 654–659. https://doi.org/10.1016/j.tins.2018.08.006

Ferreira, C. S., Charest, I., & Wimber, M. (2019). Retrieval aids the creation of a generalised memory trace and strengthens episode-unique information. *NeuroImage*, *201*. https://doi.org/10.1016/j.neuroimage.2019.07.009

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, L. E., Bowman, C. R., & Zeithamova, D. (2019). Differential Functional Connectivity along the Long Axis of the Hippocampus Aligns with Differential Role in Memory Specificity and Generalization. *Journal of Cognitive Neuroscience*, *31*(12), 1958–1975. https://doi.org/10.1162/jocn_a_01457

Frankland, P. W., & Bontempi, B. (2005). The organization of recent and remote memories. *Nature Reviews Neuroscience*, *6*(2), 119–130. https://doi.org/10.1038/nrn1607

Friston, K. J., Glaser, D. E., Henson, R. N. A., Kiebel, S., Phillips, C., & Ashburner, J. (2002). Classical and Bayesian Inference in Neuroimaging: Applications. *NeuroImage*, *16*(2), 484–512. https://doi.org/10.1006/nimg.2002.1091

Gais, S., Lucas, B., & Born, J. (2006). Sleep after learning aids memory recall. *Learning & Memory*, *13*(3), 259–262. https://doi.org/10.1101/lm.132106

Garcia-Bajos, E., Migueles, M., & Anderson, M. C. (2009). Script knowledge modulates retrieval-induced forgetting for eyewitness events. *Memory*, *17*(1), 92–103. https://doi.org/10.1080/09658210802572454

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

Golinkoff, R. M., Hirsh-Pasek, K., Bailey, L. M., & Wenger, N. R. (1992). Young children and adults use lexical principles to learn new nouns. *Developmental Psychology*, *28*(1), 99–108. https://doi.org/10.1037/0012-1649.28.1.99

Graham, K. S., Simons, J. S., Pratt, K. H., Patterson, K., & Hodges, J. R. (2000). Insights from semantic dementia on the relationship between episodic and semantic memory. *Neuropsychologia*, *38*(3), 313–324. https://doi.org/10.1016/S0028-3932(99)00073-1

Graves, K. N., Antony, J. W., & Turk-Browne, N. B. (2020). Finding the Pattern: On-Line Extraction of Spatial Structure During Virtual Navigation. *Psychological Science*, *31*(9), 1183–1190. https://doi.org/10.1177/0956797620948828

Greenberg, Daniel. L., & Verfaellie, M. (2010). Interdependence of episodic and semantic memory: Evidence from neuropsychology. *Journal of the International Neuropsychological Society*, *16*(5), 748–753. https://doi.org/10.1017/S1355617710000676

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

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*(1), 14–23. https://doi.org/10.1016/j.tics.2005.11.006

Halgren, E., Baudena, P., Clarke, J. M., Heit, G., Liégeois, C., Chauvel, P., & Musolino, A. (1995). Intracerebral potentials to rare target and distractor auditory and visual stimuli. I. Superior temporal plane and parietal lobe. *Electroencephalography and Clinical Neurophysiology*, *94*(3), 191–220. https://doi.org/10.1016/0013-4694(94)00259-N

Hampton, J. A. (1979). Polymorphous concepts in semantic memory. *Journal of Verbal Learning and Verbal Behavior*, *18*(4), 441–461. https://doi.org/10.1016/S0022-5371(79)90246-9

Harlow, I. M., & Donaldson, D. I. (2013). Source accuracy data reveal the thresholded nature of human episodic memory. *Psychonomic Bulletin & Review*, *20*(2), 318–325. https://doi.org/10.3758/s13423-012-0340-9

Hasselmo, M. E., & Stern, C. E. (2006). Mechanisms underlying working memory for novel information. *Trends in Cognitive Sciences*, *10*(11), 487–493. https://doi.org/10.1016/j.tics.2006.09.005

Hawkins, G. E., Hayes, B. K., Donkin, C., Pasqualino, M., & Newell, B. R. (2015). A Bayesian latent-mixture model analysis shows that informative samples reduce base-rate neglect. *Decision*, *2*(4), 306–318. https://doi.org/10.1037/dec0000024

Head, H., & Holmes, G. (1911). Sensory Disturbances from Cerebral Lesions. *Brain*, *34*(2–3), 102–254. https://doi.org/10.1093/brain/34.2-3.102

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

Hintzman, D. L. (1986). 'Schema abstraction' in a multiple-trace memory model. *Psychological Review*, *93*(4), 411–428. https://doi.org/10.1037/0033-295X.93.4.411

Hu, P., Stylos-Allan, M., & Walker, M. P. (2006). Sleep Facilitates Consolidation of Emotional Declarative Memory. *Psychological Science*, *17*(10), 891–898. https://doi.org/10.1111/j.1467-9280.2006.01799.x

Hunt, R. R., & Worthen, J. B. (2006). *Distinctiveness and memory*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195169669.001.0001

Incisa della Rocchetta, A., & Milner, B. (1993). Strategic search and retrieval inhibition: The role of the frontal lobes. *Neuropsychologia*, *31*(6), 503–524. https://doi.org/10.1016/0028-3932(93)90049-6

Inostroza, M., & Born, J. (2013). Sleep for Preserving and Transforming Episodic Memory. *Annual Review of Neuroscience*, *36*(1), 79–102. https://doi.org/10.1146/annurev-neuro-062012-170429

Irish, M., & Vatansever, D. (2020). Rethinking the episodic-semantic distinction from a gradient perspective. *Current Opinion in Behavioral Sciences*, *32*, 43–49. https://doi.org/10.1016/j.cobeha.2020.01.016

Jenkins, J. G., & Dallenbach, K. M. (1924). Obliviscence during Sleep and Waking. *The American Journal of Psychology*, *35*(4), 605–612. https://doi.org/10.2307/1414040

Kafkas, A., & Montaldi, D. (2014). Two separate, but interacting, neural systems for familiarity and novelty detection: A dual-route mechanism. *Hippocampus*, *24*(5), 516–527. https://doi.org/10.1002/hipo.22241

Kjelstrup, K. B., Solstad, T., Brun, V. H., Hafting, T., Leutgeb, S., Witter, M. P., Moser, E. I., & Moser, M.-B. (2008). Finite Scale of Spatial Representation in the Hippocampus. *Science*, *321*(5885), 140–143. https://doi.org/10.1126/science.1157086

Kleider, H. M., Pezdek, K., Goldinger, S. D., & Kirk, A. (2008). Schema-driven source misattribution errors: Remembering the expected from a witnessed event. *Applied Cognitive Psychology*, *22*(1), 1–20. https://doi.org/10.1002/acp.1361

Kliegl, O., Pastötter, B., & Bäuml, K.-H. T. (2015). The contribution of encoding and retrieval processes to proactive interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(6), 1778–1789. https://doi.org/10.1037/xlm0000096

Knight, R. T. (1996). Contribution of human hippocampal region to novelty detection. *Nature*, *383*(6597), 256–259. https://doi.org/10.1038/383256a0

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A Neostriatal Habit Learning System in Humans. *Science*, *273*(5280), 1399–1402. https://doi.org/10.1126/science.273.5280.1399

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Scene Memory Is More Detailed Than You Think: The Role of Categories in Visual Long-Term Memory. *Psychological Science*, *21*(11), 1551–1556. https://doi.org/10.1177/0956797610385359

Korkki, S. M., Richter, F. R., Jeyarathnarajah, P., & Simons, J. S. (2020). Healthy ageing reduces the precision of episodic memory retrieval. *Psychology and Aging*, *35*(1), 124–142. https://doi.org/10.1037/pag0000432

Korkki, S. M., Richter, F. R., & Simons, J. S. (2021). *Hippocampal-cortical encoding activity predicts the precision of episodic memory*. 2020.11.10.376814. https://doi.org/10.1101/2020.11.10.376814

Kroes, M. C. W., & Fernández, G. (2012). Dynamic neural systems enable adaptive, flexible memories. *Neuroscience & Biobehavioral Reviews*, *36*(7), 1646–1666. https://doi.org/10.1016/j.neubiorev.2012.02.014

Kumaran, D. (2013). Schema-driven facilitation of new hierarchy learning in the transitive inference paradigm. *Learning & Memory*, *20*(7), 388–394. https://doi.org/10.1101/lm.030296.113

Kumaran, D., Hassabis, D., & McClelland, J. L. (2016). What Learning Systems do Intelligent Agents Need? Complementary Learning Systems Theory Updated. *Trends in Cognitive Sciences*, *20*(7), 512–534. https://doi.org/10.1016/j.tics.2016.05.004

Kumaran, D., & Maguire, E. A. (2007). Which computational mechanisms operate in the hippocampus during novelty detection? *Hippocampus*, *17*(9), 735–748. https://doi.org/10.1002/hipo.20326

Kumaran, D., & McClelland, J. L. (2012). Generalization Through the Recurrent Interaction of Episodic Memories. *Psychological Review*, *119*(3), 573–616. https://doi.org/10.1037/a0028681

Kumaran, D., Summerfield, J. J., Hassabis, D., & Maguire, E. A. (2009). Tracking the Emergence of Conceptual Knowledge during Human Decision Making. *Neuron*, *63*(6), 889–901. https://doi.org/10.1016/j.neuron.2009.07.030

Lake, B. M., Linzen, T., & Baroni, M. (2019). Human few-shot learning of compositional instructions. *ArXiv:1901.04587 [Cs]*. http://arxiv.org/abs/1901.04587

Lampinen, J. M., Copeland, S. M., & Neuschatz, J. S. (2001). Recollections of things schematic: Room schemas revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(5), 1211–1222. https://doi.org/10.1037/0278-7393.27.5.1211

Lee, J.-H., Marzelli, M., Jolesz, F. A., & Yoo, S.-S. (2009). Automated classification of fMRI data employing trial-based imagery tasks. *Medical Image Analysis*, *13*(3), 392–404. https://doi.org/10.1016/j.media.2009.01.001

Lew, A. R., & Howe, M. L. (2017). Out of place, out of mind: Schema-driven false memory effects for object-location bindings. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(3), 404–421. https://doi.org/10.1037/xlm0000317

Liu, Z.-X., Grady, C., & Moscovitch, M. (2016). Effects of Prior-Knowledge on Brain Activation and Connectivity During Associative Memory Encoding. *Cerebral Cortex*, bhw047. https://doi.org/10.1093/cercor/bhw047

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory*, *4*(1), 19–31. https://doi.org/10.1037/0278-7393.4.1.19

Lovett, M. C., & Schunn, C. D. (1999). Task representations, strategy variability, and base-rate neglect. *Journal of Experimental Psychology: General*, *128*(2), 107–130. https://doi.org/10.1037/0096-3445.128.2.107

Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: From psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, *17*(8), 391–400. https://doi.org/10.1016/j.tics.2013.06.006

Mack, M. L., Love, B. C., & Preston, A. R. (2016). Dynamic updating of hippocampal object representations reflects new conceptual knowledge. *Proceedings of the National Academy of Sciences*, *113*(46), 13203–13208. https://doi.org/10.1073/pnas.1614048113

Mack, M. L., Love, B. C., & Preston, A. R. (2018). Building concepts one episode at a time: The hippocampus and concept formation. *Neuroscience Letters*, *680*, 31–38. https://doi.org/10.1016/j.neulet.2017.07.061

Mack, M. L., Preston, A. R., & Love, B. C. (2013). Decoding the Brain's Algorithm for Categorization from Its Neural Implementation. *Current Biology*, *23*(20), 2023–2027. https://doi.org/10.1016/j.cub.2013.08.035

Maldjian, J. A., Laurienti, P. J., Kraft, R. A., & Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage*, *19*(3), 1233–1239. https://doi.org/10.1016/S1053-8119(03)00169-1

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

Maren, S., Aharonov, G., & Fanselow, M. S. (1997). Neurotoxic lesions of the dorsal hippocampus and Pavlovian fear conditioning in rats. *Behavioural Brain Research*, *88*(2), 261–274. https://doi.org/10.1016/S0166-4328(97)00088-0

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20*(2), 121–157. https://doi.org/10.1016/0010-0285(88)90017-5

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*(3), 419–457. https://doi.org/10.1037/0033-295X.102.3.419

McKoon, G., & Ratcliff, R. (1986). Inferences about predictable events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(1), 82–91. https://doi.org/10.1037/0278-7393.12.1.82

McNab, F., & Klingberg, T. (2008). Prefrontal cortex and basal ganglia control access to working memory. *Nature Neuroscience*, *11*(1), 103–107. https://doi.org/10.1038/nn2024

Medin, D. L., Altom, M. W., & Murphy, T. D. (1984). Given versus induced category representations: Use of prototype and exemplar information in classification. *Journal* *of Experimental Psychology: Learning, Memory, and Cognition*, *10*(3), 333–352. https://doi.org/10.1037/0278-7393.10.3.333

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*(3), 207–238. https://doi.org/10.1037/0033-295X.85.3.207

Merriman, W. E., & Stevenson, C. M. (1997). Restricting a Familiar Name in Response to Learning a New One: Evidence for the Mutual Exclusivity Bias in Young Two-Year-Olds. *Child Development*, *68*(2), 211–228. https://doi.org/10.2307/1131846

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed Representations of Words and Phrases and their Compositionality. *ArXiv:1310.4546 [Cs, Stat]*. http://arxiv.org/abs/1310.4546

Mirković, J., & Gaskell, M. G. (2016). Does Sleep Improve Your Grammar? Preferential Consolidation of Arbitrary Components of New Linguistic Knowledge. *PLOS ONE*, *11*(4). https://doi.org/10.1371/journal.pone.0152489

Mirković, J., Vinals, L., & Gaskell, M. G. (2019). The role of complementary learning systems in learning and consolidation in a quasi-regular domain. *Cortex*, *116*, 228–249. https://doi.org/10.1016/j.cortex.2018.07.015

Mograss, M. A., Guillem, F., Brazzini-Poisson, V., & Godbout, R. (2009). The effects of total sleep deprivation on recognition memory processes: A study of event-related potential. *Neurobiology of Learning and Memory*, *91*(4), 343–352. https://doi.org/10.1016/j.nlm.2009.01.008

Montefinese, M., Zannino, G. D., & Ambrosini, E. (2015). Semantic similarity between old and new items produces false alarms in recognition memory. *Psychological Research*, *79*(5), 785–794. https://doi.org/10.1007/s00426-014-0615-z

Moscovitch, M., Cabeza, R., Winocur, G., & Nadel, L. (2016). Episodic Memory and Beyond: The Hippocampus and Neocortex in Transformation. *Annual Review of Psychology*, *67*, 105–134. https://doi.org/10.1146/annurev-psych-113011-143733

Moscovitch, M., & Gilboa, A. (2021). Systems consolidation, transformation and reorganization: Multiple Trace Theory, Trace Transformation Theory and their Competitors. *PsyArXiv*. https://doi.org/10.31234/osf.io/yxbrs

Murphy, G. L. (2016). Is there an exemplar theory of concepts? *Psychonomic Bulletin & Review*, *23*(4), 1035–1042. https://doi.org/10.3758/s13423-015-0834-3

Murphy, G. L., & Brownell, H. H. (1985). Category differentiation in object recognition: Typicality constraints on the basic category advantage. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*(1), 70–84. https://doi.org/10.1037/0278-7393.11.1.70

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., Samsonovich, A., Ryan, L., & Moscovitch, M. (2000). Multiple trace theory of human memory: Computational, neuroimaging, and neuropsychological results. *Hippocampus*, *10*(4), 352–368. https://doi.org/10.1002/1098-1063(2000)10:4<352::AID-HIPO2>3.0.CO;2-D

Nakamura, G. V., Graesser, A. C., Zimmerman, J. A., & Riha, J. (1985). Script processing in a natural situation. *Memory & Cognition*, *13*(2), 140–144. https://doi.org/10.3758/BF03197006

Neuschatz, J. S., Lampinen, J. M., Preston, E. L., Hawkins, E. R., & Toglia, M. P. (2002). The effect of memory schemata on memory and the phenomenological experience of naturalistic situations. *Applied Cognitive Psychology*, *16*(6), 687–708. https://doi.org/10.1002/acp.824

Nilakantan, A. S., Bridge, D. J., VanHaerents, S., & Voss, J. L. (2018). Distinguishing the precision of spatial recollection from its success: Evidence from healthy aging and unilateral mesial temporal lobe resection. *Neuropsychologia*, *119*, 101–106. https://doi.org/10.1016/j.neuropsychologia.2018.07.035

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General*, *115*(1), 39–57. https://doi.org/10.1037/0096-3445.115.1.39

Nosofsky, R. M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(4), 700–708. https://doi.org/10.1037/0278-7393.14.4.700

Nosofsky, R. M., & Zaki, S. R. (2002). Exemplar and prototype models revisited: Response strategies, selective attention, and stimulus generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(5), 924–940. https://doi.org/10.1037/0278-7393.28.5.924

Olman, C. A., Davachi, L., & Inati, S. (2009). Distortion and signal loss in medial temporal lobe. *PloS One*, *4*(12), e8160. https://doi.org/10.1371/journal.pone.0008160

O’Reilly, R. C., & Frank, M. J. (2006). Making working memory work: A computational model of learning in the prefrontal cortex and basal ganglia. *Neural Computation*, *18*(2), 283–328. https://doi.org/10.1162/089976606775093909

Packard, P. A., Rodríguez-Fornells, A., Bunzeck, N., Nicolás, B., Diego-Balaguer, R. de, & Fuentemilla, L. (2017). Semantic Congruence Accelerates the Onset of the Neural

Signals of Successful Memory Encoding. *Journal of Neuroscience*, *37*(2), 291–301. https://doi.org/10.1523/JNEUROSCI.1622-16.2016

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

Pezdek, K., Whetstone, T., Reynolds, K., Askari, N., & Dougherty, T. (1989). Memory for real-world scenes: The role of consistency with schema expectation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(4), 587–595. https://doi.org/10.1037/0278-7393.15.4.587

Piaget, J. (1926). Psychology. *The Monist*, *36*(3), 430–455.

Poppenk, J., Evensmoen, H. R., Moscovitch, M., & Nadel, L. (2013). Long-axis specialization of the human hippocampus. *Trends in Cognitive Sciences*, *17*(5), 230–240. https://doi.org/10.1016/j.tics.2013.03.005

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, *77*(3, Pt.1), 353–363. https://doi.org/10.1037/h0025953

Postman, L. (1968). Hermann Ebbinghaus. *American Psychologist*, *23*(3), 149–157. https://doi.org/10.1037/h0025659

Preston, A. R., & Eichenbaum, H. (2013). Interplay of Hippocampus and Prefrontal Cortex in Memory. *Current Biology*, *23*(17), 764–773. https://doi.org/10.1016/j.cub.2013.05.041

Preston, A. R., Molitor, R. J., Pudhiyidath, A., & Schlichting, M. L. (2017). Schemas. In J. Byrne H. (Ed.), *Learning and Memory: A Comprehensive Reference* (2nd ed., Vol. 3, pp. 125–132). Academic Press.

Preston, A. R., Shrager, Y., Dudukovic, N. M., & Gabrieli, J. D. E. (2004). Hippocampal contribution to the novel use of relational information in declarative memory. *Hippocampus*, *14*(2), 148–152. https://doi.org/10.1002/hipo.20009

Radvansky, G. A., & Tamplin, A. K. (2012). Memory. In V. S. Ramachandran (Ed.), *Encyclopedia of Human Behavior (Second Edition)* (pp. 585–592). Academic Press. https://doi.org/10.1016/B978-0-12-375000-6.00229-9

Ramasangu, H., & Sinha, N. (2014). Cognitive state classification using transformed fMRI data. *2014 International Conference on Signal Processing and Communications (SPCOM)*, 1–5. https://doi.org/10.1109/SPCOM.2014.6983987

Ranganath, C., & Rainer, G. (2003). Neural mechanisms for detecting and remembering novel events. *Nature Reviews Neuroscience*, *4*(3), 193–202. https://doi.org/10.1038/nrn1052

Rasch, B., & Born, J. (2013). About Sleep's Role in Memory. *Physiological Reviews*, *93*(2), 681–766. https://doi.org/10.1152/physrev.00032.2012

Raykov, P. P., Keidel, J. L., Oakhill, J., & Bird, C. M. (2020). 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

Reed, J. M., & Squire, L. R. (1998). Retrograde Amnesia for Facts and Events: Findings from Four New Cases. *The Journal of Neuroscience*, *18*(10), 3943–3954. https://doi.org/10.1523/JNEUROSCI.18-10-03943.1998

Richards, B. A., Xia, F., Santoro, A., Husse, J., Woodin, M. A., Josselyn, S. A., & Frankland, P. W. (2014). Patterns across multiple memories are identified over time. *Nature Neuroscience*, *17*(7), 981–986. https://doi.org/10.1038/nn.3736

Richter, F. R., Bays, P. M., Jeyarathnarajah, P., & Simons, J. S. (2019). Flexible updating of dynamic knowledge structures. *Scientific Reports*, *9*(1), Article 1. https://doi.org/10.1038/s41598-019-39468-9

Richter, F. R., Cooper, R. A., Bays, P. M., & Simons, J. S. (2016). Distinct neural mechanisms underlie the success, precision, and vividness of episodic memory. *ELife*, *5*, e18260. https://doi.org/10.7554/eLife.18260

Rips, L. J. (1975). Inductive judgments about natural categories. *Journal of Verbal Learning & Verbal Behavior*, *14*(6), 665–681. https://doi.org/10.1016/S0022-5371(75)80055-7

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

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 803–814. https://doi.org/10.1037/0278-7393.21.4.803

Rosch, E. H. (1973). Natural categories. *Cognitive Psychology*, *4*(3), 328–350. https://doi.org/10.1016/0010-0285(73)90017-0

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*(4), 573–605. https://doi.org/10.1016/0010-0285(75)90024-9

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*(3), 382–439. https://doi.org/10.1016/0010-0285(76)90013-X

Rosenbaum, R. S., Gilboa, A., Levine, B., Winocur, G., & Moscovitch, M. (2009). Amnesia as an impairment of detail generation and binding: Evidence from personal, fictional, and semantic narratives in K.C. *Neuropsychologia*, *47*(11), 2181–2187. https://doi.org/10.1016/j.neuropsychologia.2008.11.028

Rosenbaum, R. S., Priselac, S., Köhler, S., Black, S. E., Gao, F., Nadel, L., & Moscovitch, M. (2000). Remote spatial memory in an amnesic person with extensive bilateral hippocampal lesions. *Nature Neuroscience*, *3*(10), 1044–1048. https://doi.org/10.1038/79867

Sadeh, T., Ozubko, J. D., Winocur, G., & Moscovitch, M. (2016). Forgetting Patterns Differentiate Between Two Forms of Memory Representation. *Psychological Science*, *27*(6), 810–820. https://doi.org/10.1177/0956797616638307

Sakamoto, Y., & Love, B. C. (2004). Schematic Influences on Category Learning and Recognition Memory. *Journal of Experimental Psychology: General*, *133*(4), 534–553. https://doi.org/10.1037/0096-3445.133.4.534

Schlichting, M. L., Mumford, J. A., & Preston, A. R. (2015). Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. *Nature Communications*, *6*(1), 8151. https://doi.org/10.1038/ncomms9151

Schurgin, M. W., Wixted, J. T., & Brady, T. F. (2020). Psychophysical scaling reveals a unified theory of visual memory strength. *Nature Human Behaviour*, *4*(11), 1156–1172. https://doi.org/10.1038/s41562-020-00938-0

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry*, *20*(1), 11–21.

Sekeres, M. J., Bonasia, K., St-Laurent, M., Pishdadian, S., Winocur, G., Grady, C., & Moscovitch, M. (2016). Recovering and preventing loss of detailed memory: Differential rates of forgetting for detail types in episodic memory. *Learning & Memory*, *23*(2), 72–82. https://doi.org/10.1101/lm.039057.115

Sharon, T., Moscovitch, M., & Gilboa, A. (2011). Rapid neocortical acquisition of long-term arbitrary associations independent of the hippocampus. *Proceedings of the National Academy of Sciences*, *108*(3), 1146–1151. https://doi.org/10.1073/pnas.1005238108

Shiffrin, R. M., Huber, D. E., & Marinelli, K. (1995). Effects of category length and strength on familiarity in recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(2), 267–287. https://doi.org/10.1037/0278-7393.21.2.267

Shohamy, D., & Wagner, A. D. (2008). Integrating Memories in the Human Brain: Hippocampal-Midbrain Encoding of Overlapping Events. *Neuron*, *60*(2), 378–389. https://doi.org/10.1016/j.neuron.2008.09.023

Smith, D. J., & Minda, J. P. (2000). Thirty categorization results in search of a model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(1), 3–27. https://doi.org/10.1037/0278-7393.26.1.3

Spalding, K. N., Schlichting, M. L., Zeithamova, D., Preston, A. R., Tranel, D., Duff, M. C., & Warren, D. E. (2018). Ventromedial Prefrontal Cortex Is Necessary for Normal Associative Inference and Memory Integration. *The Journal of Neuroscience*, *38*(15), 3767–3775. https://doi.org/10.1523/JNEUROSCI.2501-17.2018

Spiers, H. J., Maguire, E. A., & Burgess, N. (2001). Hippocampal Amnesia. *Neurocase*, *7*(5), 357–382. https://doi.org/10.1076/neur.7.5.357.16245

Squire, L. R. (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiology of Learning and Memory*, *82*(3), 171–177. https://doi.org/10.1016/j.nlm.2004.06.005

Strange, B. A., Hurlemann, R., Duggins, A., Heinze, H.-J., & Dolan, R. J. (2005). Dissociating intentional learning from relative novelty responses in the medial temporal lobe. *NeuroImage*, *25*(1), 51–62. https://doi.org/10.1016/j.neuroimage.2004.12.014

Sun, S. Z., Fidalgo, C., Barense, M. D., Lee, A. C. H., Cant, J. S., & Ferber, S. (2017). Erasing and blurring memories: The differential impact of interference on separate aspects of forgetting. *Journal of Experimental Psychology: General*, *146*(11), 1606–1630. https://doi.org/10.1037/xge0000359

Sweegers, C. C. G., Coleman, G. A., van Poppel, E. a. M., Cox, R., & Talamini, L. M. (2015). Mental Schemas Hamper Memory Storage of Goal-Irrelevant Information. *Frontiers in Human Neuroscience*, *9*(629). https://doi.org/10.3389/fnhum.2015.00629

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

Takashima, A., Nieuwenhuis, I. L. C., Jensen, O., Talamini, L. M., Rijpkema, M., & Fernández, G. (2009). Shift from Hippocampal to Neocortical Centered Retrieval Network with Consolidation. *Journal of Neuroscience*, *29*(32), 10087–10093. https://doi.org/10.1523/JNEUROSCI.0799-09.2009

Tamminen, J., Davis, M. H., Merkx, M., & Rastle, K. (2012). The role of memory consolidation in generalisation of new linguistic information. *Cognition*, *125*(1), 107–112. https://doi.org/10.1016/j.cognition.2012.06.014

Tompary, A., & Thompson-Schill, S. L. (2021). Semantic influences on episodic memory distortions. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0001017

Tompary, A., Zhou, W., & Davachi, L. (2020). Schematic memories develop quickly, but are not expressed unless necessary. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-73952-x

Toth, J. P., & Hunt, R. R. (1999). Not one versus many, but zero versus any: Structure and function in the context of the multiple memory systems debate. In *Memory: Systems, process, or function?* (pp. 232–272). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198524069.003.0009

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. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson, *Organization of Memory* (pp. 381–402). Academic Press.

Tulving, E. (1986). What kind of a hypothesis is the distinction between episodic and semantic memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(2), 307–311. https://doi.org/10.1037/0278-7393.12.2.307

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

Tulving, E., Schacter, D. L., McLachlan, D. R., & Moscovitch, M. (1988). Priming of semantic autobiographical knowledge: A case study of retrograde amnesia. *Brain and Cognition*, *8*(1), 3–20. https://doi.org/10.1016/0278-2626(88)90035-8

Turney, I. C., & Dennis, N. A. (2017). Elucidating the neural correlates of related false memories using a systematic measure of perceptual relatedness. *NeuroImage*, *146*, 940–950. https://doi.org/10.1016/j.neuroimage.2016.09.005

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., & Joliot, M. (2002). Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain. *NeuroImage*, *15*(1), 273–289. https://doi.org/10.1006/nimg.2001.0978

Underwood, B. J. (1957). Interference and forgetting. *Psychological Review*, *64*(1), 49–60. https://doi.org/10.1037/h0044616

van Buuren, M., 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. *Journal of Neuroscience*, *34*(50), 16662–16670. https://doi.org/10.1523/JNEUROSCI.2365-14.2014

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). Subtlex-UK: A New and Improved Word Frequency Database for British English. *The Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

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., 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

Vargha-Khadem, F., Gadian, D. G., Watkins, K. E., Connelly, A., Van Paesschen, W., & Mishkin, M. (1997). Differential effects of early hippocampal pathology on episodic and semantic memory. *Science*, *277*(5324), 376–380. https://doi.org/10.1126/science.277.5324.376

Wagner, I. C., van Buuren, M., Kroes, M. C., Gutteling, T. P., van der Linden, M., Morris, R. G., & Fernández, G. (2015). Schematic memory components converge within angular gyrus during retrieval. *ELife*, *4*. https://doi.org/10.7554/eLife.09668

Wagner, U., Kashyap, N., Diekelmann, S., & Born, J. (2007). The impact of post-learning sleep vs. Wakefulness on recognition memory for faces with different facial expressions. *Neurobiology of Learning and Memory*, *87*(4), 679–687. https://doi.org/10.1016/j.nlm.2007.01.004

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. *Journal of Neuroscience*, *34*(22), 7677–7682. https://doi.org/10.1523/JNEUROSCI.0119-14.2014

Watkins, O. C., & Watkins, M. J. (1975). Buildup of proactive inhibition as a cue-overload effect. *Journal of Experimental Psychology: Human Learning and Memory*, *1*(4), 442–452. https://doi.org/10.1037/0278-7393.1.4.442

Webb, C. E., Turney, I. C., & Dennis, N. A. (2016). What's the gist? The influence of schemas on the neural correlates underlying true and false memories. *Neuropsychologia*, *93*, 61–75. https://doi.org/10.1016/j.neuropsychologia.2016.09.023

Welsh, M. B., & Navarro, D. J. (2012). Seeing is believing: Priors, trust, and base rate neglect. *Organizational Behavior and Human Decision Processes*, *119*(1), 1–14. https://doi.org/10.1016/j.obhdp.2012.04.001

Winocur, G. (1990). Anterograde and retrograde amnesia in rats with dorsal hippocampal or dorsomedial thalamic lesions. *Behavioural Brain Research*, *38*(2), 145–154. https://doi.org/10.1016/0166-4328(90)90012-4

Winocur, G., & Moscovitch, M. (2011). Memory Transformation and Systems Consolidation. *Journal of the International Neuropsychological Society*, *17*(5), 766–780. https://doi.org/10.1017/S1355617711000683

Winocur, G., Moscovitch, M., & Bontempi, B. (2010). Memory formation and long-term retention in humans and animals: Convergence towards a transformation account of hippocampal–neocortical interactions. *Neuropsychologia*, *48*(8), 2339–2356. https://doi.org/10.1016/j.neuropsychologia.2010.04.016

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

Winograd, E. (1981). Elaboration and distinctiveness in memory for faces. *Journal of Experimental Psychology: Human Learning and Memory*, *7*(3), 181–190. https://doi.org/10.1037/0278-7393.7.3.181

Wismer, A. J., & Bohil, C. J. (2017). Base-rate sensitivity through implicit learning. *PLOS ONE*, *12*(6), e0179256. https://doi.org/10.1371/journal.pone.0179256

Wixted, J. T. (2004). The Psychology and Neuroscience of Forgetting. *Annual Review of Psychology*, *55*(1), 235–269. https://doi.org/10.1146/annurev.psych.55.090902.141555

Wixted, J. T., & Rohrer, D. (1993). Proactive interference and the dynamics of free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(5), 1024–1039. https://doi.org/10.1037/0278-7393.19.5.1024

Wolfe, C. R. (2007). Adaptive redundancy, denominator neglect, and the base-rate fallacy. *Behavioral and Brain Sciences*, *30*(3), 286–287. https://doi.org/10.1017/S0140525X07001951

Yamada, R., & Itsukushima, Y. (2013). The schema provokes a disparity of false recollection between actions and objects in an everyday scene. *Scandinavian Journal of Psychology*, *54*(4), 276–282. https://doi.org/10.1111/sjop.12051

Zeithamova, D., Dominick, A. L., & Preston, A. R. (2012). Hippocampal and Ventral Medial Prefrontal Activation during Retrieval-Mediated Learning Supports Novel Inference. *Neuron*, *75*(1), 168–179. https://doi.org/10.1016/j.neuron.2012.05.010

Zeithamova, D., Maddox, W. T., & Schnyer, D. M. (2008). Dissociable Prototype Learning Systems: Evidence from Brain Imaging and Behavior. *Journal of Neuroscience*, *28*(49), 13194–13201. https://doi.org/10.1523/JNEUROSCI.2915-08.2008

Zeithamova, D., Manthuruthil, C., & Preston, A. R. (2016). Repetition suppression in the medial temporal lobe and midbrain is altered by event overlap. *Hippocampus*, *26*(11), 1464–1477. https://doi.org/10.1002/hipo.22622

Zeithamova, D., & Preston, A. R. (2010). Flexible Memories: Differential Roles for Medial Temporal Lobe and Prefrontal Cortex in Cross-Episode Binding. *Journal of Neuroscience*, *30*(44), 14676–14684. https://doi.org/10.1523/JNEUROSCI.3250-10.2010

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

Zola-Morgan, S. M., & Squire, L. R. (1990). The primate hippocampal formation: Evidence for a time-limited role in memory storage. *Science*, *250*(4978), 288–290. https://doi.org/10.1126/science.2218534